

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2004, 17:34:08 ; Search time 41 seconds  
(without alignments)

4701.593 Million cell updates/sec

Title: US-10-066-269-58

Perfect score: 4029

Sequence: 1 MGRGPDAGPRLPLLLL.....VGLSLAVVILTLFSLSAV 747

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_protein:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacterioph:\*
- 17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3300	81.9	620	Q8WX60	Q8WX60 homo sapien
2	3002	74.5	570	Q9NTA7	Q9NTA7 homo sapien
3	2957	73.4	570	Q96DM9	Q96DM9 homo sapien
4	2121	52.6	409	Q9H6V2	Q9H6V2 homo sapien
5	1792	44.5	345	Q8WU83	Q8WU83 homo sapien
6	335.5	8.3	662	11 Q9ESCO	Q9ESCO rattus norv
7	335.5	8.3	738	11 Q9ESCI	Q9ESCI rattus norv
8	331.5	8.2	738	11 Q9IW44	Q9IW44 mus musculus
9	331.5	8.2	739	11 Q923X1	Q923X1 mus musculus
10	306.5	7.6	3567	11 Q9ES77	Q9ES77 mus musculus
11	299	7.4	937	5 Q9BLJ1	Q9BLJ1 ciona intes
12	294.5	7.3	3857	11 Q98840	Q98840 mus musculus
13	291	7.2	2189	5 Q9BI05	Q9BI05 elmeria ten
14	290.5	7.2	2872	11 Q9WU88	Q9WU88 rattus norv
15	285.5	7.1	1587	4 Q90508	Q90508 homo sapien
16	284.5	7.1	1511	4 Q75412	Q75412 homo sapien

#### ALIGNMENTS

##### RESULT 1

Q8WX60 PRELIMINARY; PRT; 620 AA.

ID Q8WX60  
AC Q8WX60;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2003 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE BA01.1 (Novel protein) (Fragment).  
GN BA01.1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sehra H.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL138756; CAD13445.1; -  
DR InterPro; IPR00152; ASX\_hydroxyl.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
DR Pfam; PF00884; sushi; 2.  
DR SMART; SM00032; CCF; 2.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
KW EGF-like domain.  
FT NON TER 1  
SQ SEQUENCE 620 AA; 69174 MW; 9BESC9817E842CFF CRC64;

Query Match 81.9%; Score 3300; DB 4; Length 620;

Best Local Similarity 100.0%; Pred. No. 1.5e-276;

Matches 620; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 ECEVSLCRHGRCVNTGHSPECYCMGCLPRNGPEPFPHFTTDTSCTEIDCGTFPEVPD 187

DB 1 ECEVSLCRHGRCVNTGHSPECYCMGCLPRNGPEPFPHFTTDTSCTEIDCGTFPEVPD 60

QY 188 GVIIGNYISSLSGQVRYACRGFFSVDPEDTVSSCTGLTWESPKLHCQINCNGPPMRH 247

Q00718 homo sapien  
Q81YGO homo sapien  
Q60789 mus musculus  
Q8NG72 homo sapien  
Q9NG36 homo sapien  
Q8AXM6 xenopus lae  
Q8JFZ4 xenopus lae  
Q8K4G0 mus musculus  
Q8K4G1 mus musculus  
Q9WU88 rattus norv  
Q8GAT7 homo sapien  
Q89999 mus musculus  
Q17494 caenorhabdi  
Q8TER0 homo sapien  
Q8WPL1 oikopleura  
Q14767 homo sapien  
Q9JJS0 mus musculus  
Q96JP8 homo sapien  
P87363 gallus gall  
Q62285 mus musculus  
Q91X17 mus musculus  
Q9UHX3 homo sapien  
Q8NG96 homo sapien  
Q35806 rattus norv  
Q28019 bos taurus  
Q8SS3 dictyosteli  
Q9Y3V7 homo sapien  
Q8IUI1 homo sapien  
Q8CG18 mus musculus

```

Db 61 GYIIGNYSSLSQVRYACREFFSVPEDTVSSCTGLTWSPKHLHQEI 120
QY 248 AILVGNHSLRGVARYVCOQFESPGGKI:SVCTEKGWTRESLTTCETILTINDVSLF 307
Db 121 AILVGNHSLRGVARYVCOQFESPGGKI:SVCTEKGWTRESLTTCETILTINDVSLF 180
QY 308 NDTCVRWQNSRRINPKISVY:SIKQRLD:MEVREESTVNLTTDSRTPVCLALYPGTN 367
Db 181 NDTCVRWQNSRRINPKISVY:SIKQRLD:MEVREESTVNLTTDSRTPVCLALYPGTN 240
QY 368 YTNVISTAPPRSMPPAVIGFQTAEDVLLD:GSENFISFNCTCLKNRRSRKVGSEHYQ 427
Db 241 YTNVISTAPPRSMPPAVIGFQTAEDVLLD:GSENFISFNCTCLKNRRSRKVGSEHYQ 300
QY 428 FTVLGQWYLANFNSHATSNFTTRQVPV:LDLYPTTDY:VNVTLRSKRSVQIITA 487
Db 301 FTVLGQWYLANFNSHATSNFTTRQVPV:LDLYPTTDY:VNVTLRSKRSVQIITA 360
QY 488 TTPAVKQI:SNISGNETCLWRSIKTADM:EMVLFHWQWYKQKPAQEMFTN:SSSS 547
Db 361 TTPAVKQI:SNISGNETCLWRSIKTADM:EMVLFHWQWYKQKPAQEMFTN:SSSS 420
QY 548 RDEVECLDLRPGTNNVLSRALSSSELPVVI:3LTQITEPPLPEVEFFTVHGRPLRLRLR 607
Db 421 RDEVECLDLRPGTNNVLSRALSSSELPVVI:3LTQITEPPLPEVEFFTVHGRPLRLRLR 480
QY 608 KAKENKNGPISSQV:LVPLALQSTFSCDSE:ASFFSNASDADGYAAELAKVDPDAM 667
Db 481 KAKENKNGPISSQV:LVPLALQSTFSCDSE:ASFFSNASDADGYAAELAKVDPDAM 540
QY 668 EPIGDRLYYGEYNNAPLKRGSYDICIILRI:SEWNVKRRHS:CAVWAQVQDSSLMLLQMAQ 727
Db 541 EPIGDRLYYGEYNNAPLKRGSYDICIILRI:SEWNVKRRHS:CAVWAQVQDSSLMLLQMAQ 600
QY 728 VGLGSLAVVILITLFSFSAV 747
Db 601 VGLGSLAVVILITLFSFSAV 620

```

## RESULT 2

```

Q9NTA7 ID Q9NTA7 PRELIMINARY; PRT; 570 AA.
AC Q9NTA7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP761E1824.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amalgam;
RA Bloembergen H., Boecher M., Brandt P., Mewes H.W., Gassenhuber J.,
RA Wilmanns S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137432; CAB70735.1; -.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00084; sushi; 2.
DR SMART; SM00032; CCP; 2.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 570 AA; 63713 MW; 66B1A4B4096C3A1B CRC64;

```

Query Match 74.5%; Score 3002; DB 4; Length 570;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-251;  
 Matches 570; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 DCGTPEPPDGYIIGNYSSLSQVRYACREFFSVPEDTVSSCTGLTWSPKHLHQEI 237

```

Db 1 DCGTPEPPDGYIIGNYSSLSQVRYACREFFSVPEDTVSSCTGLTWSPKHLHQEI 60
QY 238 NCGNPPPMRRAILVGNHSLRGVARYVCOQFESPGGKI:SVCTEKGWTRESLTTCETILTINDVSLF 297
Db 61 NCGNPPPMRRAILVGNHSLRGVARYVCOQFESPGGKI:SVCTEKGWTRESLTTCETILTINDVSLF 120
QY 298 LTKINDVSLNDCTCVRWQNSRRINPKISVY:SIKQRLD:MEVREESTVNLTTDSRTPVCLALYPGTN 357
Db 121 LTKINDVSLNDCTCVRWQNSRRINPKISVY:SIKQRLD:MEVREESTVNLTTDSRTPVCLALYPGTN 180
QY 358 VCLALYPGTNNVISTAPPRSMPPAVIGFQTAEDVLLD:GSENFISFNCTCLKNRRSRKVGSEHYQ 417
Db 181 VCLALYPGTNNVISTAPPRSMPPAVIGFQTAEDVLLD:GSENFISFNCTCLKNRRSRKVGSEHYQ 240
QY 418 RKVGESEHYQ:FTVLGQWYLANFNSHATSNFTTRQVPV:LDLYPTTDY:VNVTLRSKRSVQIITA 477
Db 241 RKVGESEHYQ:FTVLGQWYLANFNSHATSNFTTRQVPV:LDLYPTTDY:VNVTLRSKRSVQIITA 300
QY 478 KHSVQIITATTPAVKQI:SNISGNETCLWRSIKTADM:EMVLFHWQWYKQKPAQEMFTN:SSSS 537
Db 301 KHSVQIITATTPAVKQI:SNISGNETCLWRSIKTADM:EMVLFHWQWYKQKPAQEMFTN:SSSS 420
QY 538 EMFTN:SSSSRDPEVCLDLRPGTNNVLSRALSSSELPVVI:3LTQITEPPLPEVEFFTVHGRPLRLRLR 597
Db 361 EMFTN:SSSSRDPEVCLDLRPGTNNVLSRALSSSELPVVI:3LTQITEPPLPEVEFFTVHGRPLRLRLR 480
QY 598 RGLPLRLRLRKAKENKNGPISSQV:LVPLALQSTFSCDSE:ASFFSNASDADGYAAELAKVDPDAM 657
Db 421 RGLPLRLRLRKAKENKNGPISSQV:LVPLALQSTFSCDSE:ASFFSNASDADGYAAELAKVDPDAM 540
QY 658 LAKVDPDAM:EIPIGDRLYYGEYNNAPLKRGSYDICIILRI:SEWNVKRRHS:CAVWAQVQDSSLMLLQMAQ 717
Db 481 LAKVDPDAM:EIPIGDRLYYGEYNNAPLKRGSYDICIILRI:SEWNVKRRHS:CAVWAQVQDSSLMLLQMAQ 600
QY 718 VGLGSLAVVILITLFSFSAV 747
Db 541 VGLGSLAVVILITLFSFSAV 620

```

## RESULT 3

```

Q96DM9 ID Q96DM9 PRELIMINARY; PRT; 570 AA.
AC Q96DM9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ32142 (C4BP).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Oshima A., Takahashi-Pujii A., Tanase T., Inose N., Takeuchi K.,
RA Arita M., Muraashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otaki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK056704; BAB71259.1; -.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00084; sushi; 2.
DR SMART; SM00032; CCP; 2.
KW Hypothetical protein.
SQ SEQUENCE 570 AA; 64211 MW; E67489C3A38E868A CRC64;

```

Query Match 73.4%; Score 2957; DB 4; Length 570;  
 Best Local Similarity 98.9%; Pred. No. 6.1e-247;  
 Matches 554; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 153 MDGYPNGPEPPPTDTSCTEIDCGTPEVDPDGYIIGNYTSISGQVRVACREGPFS 212  
DB 1 MDGYPNGPEPPPTDTSCTEIDCGTPEVDPDGYIIGNYTSISGQVRVACREGPFS 60  
QY 213 VPEDTVSSCTGLGTWSPKHCQECINCGNPPENRHAIVGNHSSRLGGVARYVCOEGPES 272  
DB 61 VPEDTVSSCTGLGTWSPKHCQECINCGNPPENRHAIVGNHSSRLGGVARYVCOEGPES 120  
QY 273 PGGKITSVCTEKGWRETSITCTEILTKINDVSLFNDTCVRQWINSRRINPKISVISIK 332  
DB 121 PGGKITSVCTEKGWRETSITCTEILTKINDVSLFNDTCVRQWINSRRINPKISVISIK 180  
QY 333 GQRLDPMSVREBTNLTDSRTPEVCLALYPGTNTVNNISAPPRRSMPAVIGFQTAEV 392  
DB 181 GQRLDPMSVREBTNLTDSRTPEVCLALYPGTNTVNNISAPPRRSMPAVIGFQTAEV 240  
QY 393 DLLEDDGGSFNISFNCTCLNRRSRKVGSEHWYQTVLGQWYLANFSHATSFNFTTRE 452  
DB 241 DLLEDDGGSFNISFNCTCLNRRSRKVGSEHWYQTVLGQWYLANFSHATSFNFTTRE 300  
QY 453 QVPVCLDLPTDYTYNVNLLSPKHSVC/TIATPPAVKQTIISNIGFNETCLRWRSI 512  
DB 301 QVPVCLDLPTDYTYNVNLLSPKHSVC/TIATPPAVKQTIISNIGFNETCLRWRSI 360  
QY 513 KTADMEEMYLPHIWGORWYQKEFAQEMTFN1SSSRDPEVCLDLPRTNYSVLRALSSE 572  
DB 361 KTADMEEMYLPHIWGORWYQKEFAQEMTFN1SSSRDPEVCLDLPRTNYSVLRALSSE 420  
QY 573 LPVVISLTTQITPPPEVEFFVTHRGPLRLRKAKEKNGPISSYQVVLPLALQSTF 632  
DB 421 LPVVISLTTQITPPPEVEFFVTHRGPLRLRKAKEKNGPISSYQVVLPLALQSTF 480  
QY 633 SCDSGASSPFSNASDADGVAAELLAKDVPIDAMEIPIGDRLYYGEYNAFLKRGSDYC 692  
DB 481 SCDSGASSPFSNASDADGVAAELLAKDVPIDAMEIPIGDRLYYGEYNAFLKRGSDYC 540  
QY 693 IILRITSEWKNRRHSCAVW 712  
DB 541 IILRITSEWKNRRHSCAVW 560

## RESULT 4

Q9H5V2 Q9H5V2 PRELIMINARY; PRT; 409 AA.  
AC Q9H5V2; SEQUENCE FROM N.A.  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
DE Hypothetical protein FLJ21833.  
OS Homo sapiens (human)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
CX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,  
RA Tanaka T., Nakamura Y., Isegaki T., Sugano S.,  
RT "NEDO human cDNA sequencing project."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK025486; BAB51549.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 409 AA; B59E76A0D1A10DA CRC64;  
Query Match 52.6%; Score 2121; DB 4; Length 409;  
Best Local Similarity 100.0%; Pred. No. 7,8e-175; Indels 0; Gaps 0;  
Matches 409; Conservative 0; Mismatches 0;  
QY 339 MESVREBTNLTDSRTPEVCLALYPGTNTVNNISAPPRRSMPAVIGFQTAEVLLLEDD 398  
DB 1 MESVREBTNLTDSRTPEVCLALYPGTNTVNNISAPPRRSMPAVIGFQTAEVLLLEDD 60

QY 399 GSENIPIFNETCLNRRSRKVGSEHWYQTVLGQWYLANFSHATSFNFTTREQVPPVC 458  
DB 61 GSENIPIFNETCLNRRSRKVGSEHWYQTVLGQWYLANFSHATSFNFTTREQVPPVC 120  
QY 459 LDLYPTDVTYVNTLLRSPKHSVC/TIATPPAVKQTIISNIGFNETCLRWRSIKTADME 518  
DB 121 LDLYPTDVTYVNTLLRSPKHSVC/TIATPPAVKQTIISNIGFNETCLRWRSIKTADME 180  
QY 519 EMYLPHIWGORWYQKEFAQEMTFN1SSSRDPEVCLDLPRTNYSVLRALSSELPVVIS 578  
DB 181 EMYLPHIWGORWYQKEFAQEMTFN1SSSRDPEVCLDLPRTNYSVLRALSSELPVVIS 240  
QY 579 LTTQITTEPPLPEVEFFVTHRGPLRLRKAKEKNGPISSYQVVLPLALQSTFSCDSEG 638  
DB 241 LTTQITTEPPLPEVEFFVTHRGPLRLRKAKEKNGPISSYQVVLPLALQSTFSCDSEG 300  
QY 639 ASFFSNASDADGVAAELLAKDVPIDAMEIPIGDRLYYGEYNAFLKRGSDYCIILRIT 698  
DB 301 ASFFSNASDADGVAAELLAKDVPIDAMEIPIGDRLYYGEYNAFLKRGSDYCIILRIT 360  
QY 699 SEWKNRRHSCAVWQKDSLSMLLQAGVGLGSLAVVILITFLSFSAV 747  
DB 361 SEWKNRRHSCAVWQKDSLSMLLQAGVGLGSLAVVILITFLSFSAV 409

## RESULT 5

Q8WU83 Q8WU83 PRELIMINARY; PRT; 345 AA.  
AC Q8WU83; SEQUENCE FROM N.A.  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
DE Hypothetical protein (Fragment).  
OS Homo sapiens (human)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
CX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RA Strausberg R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC021125; AAH21125.1; -  
KW Hypothetical protein.  
FT NON-TER 1  
SQ SEQUENCE 345 AA; 39009 MW; EBAIE3711BB95B9 CRC64;  
Query Match 44.5%; Score 1792; DB 4; Length 345;  
Best Local Similarity 100.0%; Pred. No. 1.8e-146; Indels 0; Gaps 0;  
Matches 345; Conservative 0; Mismatches 0;  
QY 403 ISIFNETCLNRRSRKVGSEHWYQTVLGQWYLANFSHATSFNFTTREQVPPVCCLDLY 462  
DB 1 ISIFNETCLNRRSRKVGSEHWYQTVLGQWYLANFSHATSFNFTTREQVPPVCCLDLY 60  
QY 463 PITDVTYVNTLLRSPKHSVC/TIATPPAVKQTIISNIGFNETCLRWRSIKTADMEEMYL 522  
DB 61 PITDVTYVNTLLRSPKHSVC/TIATPPAVKQTIISNIGFNETCLRWRSIKTADMEEMYL 120  
QY 523 FHIWGORWYQKEFAQEMTFN1SSSRDPEVCLDLPRTNYSVLRALSSELPVVISLTTQ 582  
DB 121 FHIWGORWYQKEFAQEMTFN1SSSRDPEVCLDLPRTNYSVLRALSSELPVVISLTTQ 180  
QY 583 ITEPPLPEVEFFVTHRGPLRLRKAKEKNGPISSYQVVLPLALQSTFSCDSEGASSF 642  
DB 181 ITEPPLPEVEFFVTHRGPLRLRKAKEKNGPISSYQVVLPLALQSTFSCDSEGASSF 240  
QY 643 FSNASDADGVAAELLAKDVPIDAMEIPIGDRLYYGEYNAFLKRGSDYCIILRITSEWN 702  
DB 241 FSNASDADGVAAELLAKDVPIDAMEIPIGDRLYYGEYNAFLKRGSDYCIILRITSEWN 300  
QY 703 KYRRHSCAVWQKDSLSMLLQAGVGLGSLAVVILITFLSFSAV 747  
DB 303 KYRRHSCAVWQKDSLSMLLQAGVGLGSLAVVILITFLSFSAV 409







Query Match 7.2%; Score 291; DB 5; Length 2189;  
Best Local Similarity 30.8%; Pred.No.1.9e-15;  
Matches 94; Conservative 27; Mismatches 80; Indels 104; Gaps 22;  
QY 36 LDVCA-----TCHETATCOQREKKIKICNTGVVNGRGTQVDKNCQFATLVCGHTSC 91  
Db 600 IDECANGTHCHASATCINTGQSPCAACAGSGNG-VECNVDVSCSDAD--DCCENTLC 657

Query Match	7.2%;	Score	290.5;	DB	11;	Length	2872;
Best Local Similarity	25.9%;	Pred. No.	3.2e-15;				
Matches	78;	Conservative	36;	Mismatches	105;	Indels	81;
Gaps	12;						

  

QY	37	DVATCHEATQQRGGKICICNYGVFG-NQRTQCVKNEQFGATLVCGNHTSCHNTP	95
	:		:
	:		:
	:		:
	:		:
DB	1290	NICLS----GTCENTKGFICHDGMGYSKKGKGTGTDINECEIGAH-NCGRHAVCTNTA	1344

  

QV	96	GGPVCICIGRYATNNKTFIPNDGTFCTDIDEC-EVSGLCRGGRCVNWTHGFEYCYCMD	154
----	----	--	-----

Matches	78;	Conservative	36;	Mismatches	106;	Indels	81;	Gaps	12.
Qy	37	DVATCHEATQORBGKKIKICINYGFGVNGRTQCVKNECFGATLVCNGHNSCHNTP	95						
	:	:	:	:	:	:	:	:	:
	:	:	:	:	:	:	:	:	:
Db	1290	NICLS---GTCENTKSGSFICHDMDYSGKKGTGCTDINECEIGAH-NCGRHAVCTNTA	1344						
Qy	96	GGPVCICLRYGRATNNKKEFIPNDGTFCTDIDBC-EVSGLCRGGRCVNTHTGSEFYCYMD	154						

```
Db 1345 GSFKSCSPGW-----IGDKIKTD--DECSNGTHMCSQHADCKNTMGSYRCLXD 1394
QY 155 GYLPRNGPEPHPTDTSCTEIDCGTPE /PDGYIIGNYTSLSGSQVRYACREGFFSYV 214
Db 1395 GF-----TGGTCTDLD-----E:SENLNLSGNGQCLNAPAGTRCEDMGFFV 1438
QY 215 EDTVSSCTGLGTWESPKLHCOEINCGNPPEMRHAILVGNHSSRLGGVARYVCOEGFE--S 272
Db 1439 SADGRACEDINECSLPNT-CVFGTCHNLP-----GLFRCECEIGYELDR 1481
QY 273 PGKKTTSVTEKGTWRBETLTCTELLTKINDVSLFNDTCVR-----WQINSRR 320
Db 1482 SGG-----NCTDVNECLDFTTICISGNCVNTPGSYTCDPDPFELNPT 1524
QY 321 I 321
Db 1525 V 1525

RESULT 15
ID C00508 PRELIMINARY; PRT; 1587 AA.
AC C00508;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Latent TGF-beta binding protein-4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97415399; PubMed=92711198;
RA Giltay R., Kostka G., Timpl R.;
RT "Sequence and expression of a novel member (LTBP-4) of the family of
RL FEBS Lett. 411:164-168(1997).
DR EMBL; Y13622; CAA73944.1; -.
DR HSSP; P35555; 1EMN.
DR Genew; HGNC:6717; LTBP4.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002212; Fibrin-assoc.
DR InterPro; IPR003018; GAF.
DR Pfam; PF00008; EGF; 17.
DR Pfam; PF00683; TB; 4.
DR SMART; SM00179; EGF_CA; 15.
DR SMART; SM00065; GAF; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 14.
DR PROSITE; PS00022; EGF 1; 2.
DR PROSITE; PS01186; EGF 2; 12.
DR PROSITE; PS01187; EGF_CA; 17.
KW EGF-like domain.
SQ SEQUENCE 1587 AA; 169620 MW; 57A832B95FA0AE46 CRC64;
```

```
Query Match 7.1%; Score 285.5; DB 4; Length 1587;
Best Local Similarity 22.5%; Pred. No. 3.6e-15;
Matches 127; Conservative 50; Mismatches 180; Indels 207; Gaps 25;

QY 22 GLARGAAGAPGPDGLDVCATCHEAT---CQREGKKICICNYGFVN----- 66
Db 703 GFRSGPGAPQCD-VDECARSPFTCYGRQNTGSSFCQCPMGFQPNTAGSECEVDDEC 761
QY 67 -----GATQVDKNECQFQATLVCGNHTSCHNTPGG 97
Db 762 ENHLACPGQECVNSPGSQRTCPSGHHLHGRCTDVEDCSSGAP-PCGPHGCHTNTGSS 820
QY 98 FYCICLEGYRATNN-----NKTPIPDGTF----- 122
Db 821 FRCSAPGYRAPSGRPGPCADVNECLEGDFLPHGECLNTDGSFACTCAPGYRPGRGAS 880
```

```
QY 123 CTDIDCEVSGLCRHGRCVNTHGSFECVCMGYLPRNGPFPFHTTDTATSCETID---- 178
Db 881 CLDVEDCEBEDLCQ-SGICTNTDGSFECICPPGH--RAGP-----DLASCLDVDECRE 930
QY 179 -----CGTPEVPDGYIIGNYTSLSGQ--VRYACREGFFSVPE---DTVSSCTGLGTWE 228
Db 931 RGPALCGSQ-----RCENSPGSVRCVR-DCDPGYHAGPEGTCDVDDECOEYG--- 976
QY 229 SPKLHQEINCGNPP-----EMRHAILVGNHS--SRLOGV 261
Db 977 -PEI-CGAQRCEINTPGSYRCTPACDPGYQPTBGGGCDVDCECNRSFCGAHAVCONLPGS 1034
QY 262 ARYVCOEGPESGKITSVTEKGTWRBETLTCTELLTKINDVSLFNDTCVRWQINSERI 321
Db 1035 FOCLDQGYE--GARDGRCVD-----VNECETLQGV----- 1065
QY 322 NPKISYVISIKQORLDPMESVREETVNLTTDSRTPEVCLALYPGNTYTNISTAPPERSM 381
Db 1066 --GAALCENVEGSLCVCNPSPEEPDMTGRVCVPRTSVGMSPGSQPOAPVSPVLPARPP 1123
QY 382 PAVIG-----FQTAEDVLDDEGGSFNISINETCLKN-----R 415
Db 1124 PPPLSRPRKPRKPGVSGRCRCYFDTAADDACDNLILARNVT-WOECCTVGEWGSGR 1182
QY 416 RSRKVGSEHM-YQFTVLGQRYLA 438
Db 1183 IQQCPGTETAAYQSLCPHGRGYLA 1206
```

Search completed: February 6, 2004, 17:39:04  
Job time : 45 secs



GenCore version 5.1.6  
Copyright (C) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2004, 17:25:33 ; Search time 17 Seconds  
(without alignments)  
2056.409 Million cell updates/sec

Title: US-10-066-269-58  
Perfect score: 4029  
Sequence: 1 MGRGPDAGPSRRLLPLLLL.....VLGLSLAVVILITLSPSAV 747

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt 41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	294.5	7.3	2871	FNBI_MOUSE	Q61554 mus musculus
2	290.5	7.2	2871	FNBI_HUMAN	P35555 homo sapien
3	287	7.1	835	CD97_HUMAN	P48960 homo sapien
4	283.5	7.0	2871	FNBI_BOVIN	P98133 bos taurus
5	281	7.0	2907	FNBI_MOUSE	Q61555 mus musculus
6	278.5	6.9	2911	FNBI_HUMAN	P35556 homo sapien
7	277.5	6.9	640	UROM_HUMAN	P07911 homo sapien
8	274.5	6.8	866	EMBL_HUMAN	Q14246 homo sapien
9	270	6.7	2871	FNBI_PIG	Q9CV36 sus scrofa
10	260.5	6.5	1394	LTBS_HUMAN	P22064 homo sapien
11	260.5	6.5	1595	LTBL_HUMAN	Q14766 homo sapien
12	260	6.5	931	EMRI_MOUSE	Q61549 mus musculus
13	256	6.4	1389	FNBI_MOUSE	Q8CG18 mus musculus
14	256	6.4	1713	LTBL_MOUSE	Q8CG19 mus musculus
15	251	6.2	1712	LTBL_RAT	Q00918 rattus norv
16	246.5	6.1	1247	NIDO_HUMAN	P14543 homo sapien
17	246	6.1	2319	NTC3_RAT	Q9R172 rattus norv
18	245.5	6.1	643	UROM_BOVIN	P48733 bos taurus
19	245.5	6.1	703	FNBI_HUMAN	P23142 homo sapien
20	242.5	6.0	816	NEL2_MOUSE	Q61220 mus musculus
21	241.5	6.0	705	FNBI_MOUSE	Q08879 mus musculus
22	241.5	6.0	2318	NTC3_MOUSE	Q61982 mus musculus
23	240	6.0	1221	FNBI_MOUSE	P37889 mus musculus
24	238.5	5.9	816	NEL2_RAT	Q52918 rattus norv
25	236.5	5.9	704	FNBI_CHICK	Q53775 gallus gall
26	236	5.9	493	FNBI_HUMAN	Q12805 homo sapien
27	235	5.8	644	UROM_RAT	P27590 rattus norv
28	234	5.8	2201	TENA_HUMAN	P24821 homo sapien
29	233	5.8	570	FBP3_STRPU	P49013 strongyloce
30	233	5.8	2003	NTC4_HUMAN	Q99466 homo sapien
31	231.5	5.7	2321	NTC3_HUMAN	Q9UM47 homo sapien
32	231	5.7	1245	NIDO_MOUSE	P10493 mus musculus
33	231	5.7	1364	NTC4_MOUSE	P31695 mus musculus

ALIGNMENTS

RESULT 1

ID	FNBI_MOUSE	STANDARD;	PRT; 2871 AA.
AC	Q61554; Q60826;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Fibrillin-1 precursor.		
GN	FNBI OR FNBI-1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1] _		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=55130561; PubMed=7829516;		
RA	Yin W., Germiller J., Sanguinetti C., Smiley E., Pangillan T.,		
RA	Pereira L., Ramirez F., Bonadio J.;		
RT	"Primary structure and developmental expression of Fbn-1, the mouse		
RT	fibrillin gene."		
RL	J. Biol. Chem. 270:1798-1806(1995).		
RN	[2] _		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CD-1; TISSUE=Kidney;		
RA	Ota K., Kumar A., Wada J., Liu Z., Kanwar Y.S.;		
RL	Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.		
CC	!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS		
CC	THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE		
CC	LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.		
CC	!- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER		
CC	FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE		
CC	MICROFIBRILS (BY SIMILARITY).		
CC	!- SIMILARITY: Contains 47 EGF-like domains.		
CC	!- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBP) domains.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
DR	EMBL; L29454; AAA56840.1; -		
DR	EMBL; U22493; AAA64217.1; -		
DR	PIR; A55624; A55624.		
DR	HSSP; P35555; 1AEJ.		
DR	MGI; 95489; Fbn1.		
DR	InterPro; IPR000152; Asx_hydroxyl.		
DR	InterPro; IPR001881; EGF_Ca.		
DR	InterPro; IPR006209; EGF_like.		
DR	InterPro; IPR002212; Fibrill-assoc.		
DR	Pfam; PF00008; EGF; 46.		
DR	Pfam; PF00683; TB; 9.		
DR	SMART; SM00179; EGF_CA; 42.		
DR	PROSITE; PS00010; ASX_HYDROXYL; 43.		

P21783 xenopus lae  
O35568 rattus norv  
P07522 rattus norv  
O88322 mus musculu  
O99435 homo sapien  
Q09827 gallus gall  
P80095 homo sapien  
P07207 drosophila  
O77469 caenorhabdi  
P10079 strongyloce  
Q90Y54 brachydanio  
Q14112 homo sapien

34 230 5.7 2524 1 NOTC\_XENLA  
35 493 1 FBL3\_RAT  
36 229 5.7 1133 1 EGF\_RAT  
37 229 5.7 1403 1 NID2\_MOUSE  
38 228.5 5.7 816 1 NEL2\_HUMAN  
39 228.5 5.7 816 1 NEL\_CHICK  
40 225 5.6 1184 1 FBL2\_HUMAN  
41 222 5.5 2703 1 NOTC\_DROME  
42 221 5.5 798 1 FBL1\_CAEEL  
43 220.5 5.5 1064 1 FBP1\_STRPU  
44 220.5 5.5 1213 1 JAG3\_BRARE  
45 220.5 5.5 1375 1 NID2\_HUMAN



RESULT 2	FBNI_HUMAN	STANDARD;	PRT; 2871 AA.
ID	FBNI_HUMAN		
AC	P35555;		
DT	01-JUN-1994 (Rel. 29, Created)		
DT	01-JUN-1994 (Rel. 29, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Fibrillin-1 precursor.		
GN	FBNI OR FBN.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
NCBI_Taxid=9606;			
OX	[1]		
RN	SEQUENCE FROM N.A.		
RP	TISSUE=Placenta;		
RC	MEDLINE=93372860; PubMed=8364578;		
RX	Persira L.V., D'Alessio M., Ramirez F., Lynch J.R., Sykes B.,		
RA	Pangillan T., Bonadio J.;		
RA	"Genomic organization of the sequence coding for fibrillin, the		
RT	defective gene product in Marfan syndrome.";		
RL	Hum. Mol. Genet. 2:961-968(1993).		
RN	[2]		
RP	SEQUENCE OF 1-932 FROM N.A.		
RC	TISSUE=Fibroblast, and Placenta;		
RC	MEDLINE=94010947; PubMed=7691719;		
RX	Corson G.M., Chaberg S.C., Dietz H.C., Charbonneau N.L., Sakai L.Y.;		
RA	"Fibrillin binds calcium and is coded by cDNAs that reveal a		
RT	multidomain structure and alternatively spliced exons at the 5'		
RT	end.";		
RL	Genomics 17:476-484(1993).		
RN	[3]		
RP	SEQUENCE OF 899-2871 FROM N.A.		
RC	MEDLINE=91304568; PubMed=1852207;		
RX	Maslen C.L., Corson G.M., Maddox B.K., Gianville R.W., Sakai L.Y.;		
RA	"Partial cloning of a candidate gene for the Marfan syndrome.";		
RT	Nature 352:334-337(1991).		
RL	[4]		
RN	SEQUENCE OF 813-1313 FROM N.A.		
RP	MEDLINE=91304567; PubMed=1852206;		
RX	Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Saifrazi M.,		
RA	Tsipouras P., Ramirez F., Hollister D.W.;		
RT	"Linkage of Marfan syndrome and a phenotypically related disorder to		
RT			



receptor superfamily.";  
 Genomics 32:144-147(1996).  
 (3)  
 REVISIONS.  
 Hamam J.; Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: COULD BE A RECEPTOR POTENTIALLY INVOLVED IN BOTH  
 ADHESION AND SIGNALING PROCESSES EARLY AFTER LEUKOCYTE ACTIVATION.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: Contains 5 EGF-like domains.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -1- SIMILARITY: Contains 1 GPS domain.  
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD97 entry;  
 WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd97.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X84700; CAA59173.1; -  
 DR EMBL; X94630; CAA64333.1; -  
 DR EMBL; X94631; CAA64333.1; JOINED.  
 DR EMBL; X94632; CAA64333.1; JOINED.  
 DR EMBL; X94633; CAA64333.1; JOINED.  
 DR EMBL; Z99830; CAA64333.1; JOINED.  
 DR EMBL; Z99831; CAA64333.1; JOINED.  
 DR EMBL; X94634; CAA64333.1; JOINED.  
 DR EMBL; X94635; CAA64333.1; JOINED.  
 DR EMBL; X94636; CAA64333.1; JOINED.  
 DR EMBL; X94637; CAA64333.1; JOINED.  
 DR EMBL; X94638; CAA64333.1; JOINED.  
 DR EMBL; X94639; CAA64333.1; JOINED.  
 DR EMBL; X94640; CAA64333.1; JOINED.  
 DR EMBL; X94641; CAA64333.1; JOINED.  
 DR EMBL; X94642; CAA64333.1; JOINED.  
 DR EMBL; X94643; CAA64333.1; JOINED.  
 DR EMBL; X94644; CAA64333.1; JOINED.  
 DR EMBL; X94645; CAA64333.1; JOINED.  
 DR EMBL; X94646; CAA64333.1; JOINED.  
 DR EMBL; X94647; CAA64333.1; JOINED.  
 DR PIR; I37225; I37225.  
 DR HSP; P16109; IFSB.  
 DR Genew; HGNC:1711; CD97.  
 DR MIM; 601211; -  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0004930; F:G-protein coupled receptor activity; TAS.  
 DR GO; GO:0007155; P:cell adhesion; TAS.  
 DR GO; GO:0006928; P:cell motility; TAS.  
 DR GO; GO:0007267; P:cell-cell signaling; TAS.  
 DR GO; GO:0006955; P:immune response; TAS.  
 DR GO; GO:0006954; P:inflammatory response; TAS.  
 DR InterPro; IPR00152; Asx hydroxyl.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR006209; EGF\_Like.  
 DR InterPro; IPR008332; GPCR\_secretin.  
 DR InterPro; IPR000203; PKD\_cys\_rich.  
 DR Pfam; PF00002; 7tm\_2; 1.  
 DR Pfam; PF00008; EGF; 3.  
 DR Pfam; PF01825; GPS; 1.  
 DR PRINTS; PR00249; GPCRSECRETIN.  
 DR SMART; SM00179; EGF\_CA; 4.  
 DR SMART; SM00303; GPS; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
 DR PROSITE; PS01187; EGF\_CA; 4.  
 DR PROSITE; PS00221; GPS; 1.  
 DR PROSITE; PS00650; G-PROTEIN RECEPTOR F2 2; 1.  
 DR PROSITE; PS00261; G-PROTEIN RECEPTOR F2 4; 1.  
 KW Cell adhesion; Receptor; G-protein coupled receptor; Repeat; Signal;  
 Transmembrane; EGF-like domain; Glycoprotein.

FT SIGNAL	1	20	POTENTIAL.
FT CHAIN	21	835	LEUCOCYTE ANTIGEN CD97.
FT DOMAIN	21	552	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	553	572	POTENTIAL.
FT DOMAIN	573	581	CYTOPLASMIC (POTENTIAL).
FT TRANSMEM	582	601	POTENTIAL.
FT DOMAIN	602	620	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	621	642	POTENTIAL.
FT DOMAIN	643	653	CYTOPLASMIC (POTENTIAL).
FT TRANSMEM	654	674	POTENTIAL.
FT DOMAIN	675	691	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	692	712	POTENTIAL.
FT DOMAIN	713	739	CYTOPLASMIC (POTENTIAL).
FT TRANSMEM	740	760	POTENTIAL.
FT DOMAIN	761	766	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	767	789	POTENTIAL.
FT DOMAIN	790	835	CYTOPLASMIC (POTENTIAL).
FT TRANSMEM	836	835	EGF-LIKE 1.
FT DOMAIN	836	115	EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT TRANSMEM	116	159	EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN	160	208	EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT TRANSMEM	209	257	EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN	258	542	GPS.
FT TRANSMEM	492	542	CELL ATTACHMENT SITE (POTENTIAL).
FT DOMAIN	367	369	BY SIMILARITY.
FT TRANSMEM	26	36	BY SIMILARITY.
FT DOMAIN	30	42	BY SIMILARITY.
FT TRANSMEM	44	62	BY SIMILARITY.
FT DOMAIN	68	82	BY SIMILARITY.
FT TRANSMEM	76	91	BY SIMILARITY.
FT DOMAIN	93	114	BY SIMILARITY.
FT TRANSMEM	120	133	BY SIMILARITY.
FT DOMAIN	127	142	BY SIMILARITY.
FT TRANSMEM	144	158	BY SIMILARITY.
FT DOMAIN	164	177	BY SIMILARITY.
FT TRANSMEM	171	186	BY SIMILARITY.
FT DOMAIN	188	207	BY SIMILARITY.
FT TRANSMEM	213	226	BY SIMILARITY.
FT DOMAIN	220	235	BY SIMILARITY.
FT TRANSMEM	237	256	BY SIMILARITY.
FT DOMAIN	33	33	N-LINKED (GLCNAC. .) (POTENTIAL).
FT TRANSMEM	38	38	N-LINKED (GLCNAC. .) (POTENTIAL).
FT DOMAIN	108	108	N-LINKED (GLCNAC. .) (POTENTIAL).
FT TRANSMEM	203	203	N-LINKED (GLCNAC. .) (POTENTIAL).
FT DOMAIN	371	371	N-LINKED (GLCNAC. .) (POTENTIAL).
FT TRANSMEM	406	406	N-LINKED (GLCNAC. .) (POTENTIAL).
FT DOMAIN	413	413	N-LINKED (GLCNAC. .) (POTENTIAL).
FT TRANSMEM	453	453	N-LINKED (GLCNAC. .) (POTENTIAL).
FT DOMAIN	520	520	N-LINKED (GLCNAC. .) (POTENTIAL).
FT TRANSMEM	520	520	N-LINKED (GLCNAC. .) (POTENTIAL).
FT DOMAIN	121	213	MISSING (IN REF. 1).
FT TRANSMEM	835	91941	MM; 050672E7A4C12A9E CRC64;
FT DOMAIN	835	91941	MM; 050672E7A4C12A9E CRC64;
FT TRANSMEM	835	91941	MM; 050672E7A4C12A9E CRC64;

Query Match	7.1%	Score 287;	DB 1;	Length 835;
Best Local Similarity	21.8%	Pred. No. 7.1e-12;		
Matches 168;	Conservative 89;	Mismatches 242;	Indels 270;	Gaps 40;
QY	29	GAPGPDLDVCACTHEATCOOREGKCI CNYGF-----VNGRTQCVKNECQFAT	82	
Db	17	CAETQDSRGCARWCPCQNSCVNATA---CRCNPGFSFSEIITTTCTDDINECATPSK	73	
QY	83	LVCNHTSCHNTPGGYCICLEGVATNNKTFIPNDGTCTDIDCEVS-GLCRHGRC	141	
Db	74	VSCGFSDCWNTGSDGSPGPGYFVSGAKTFKNESENTQDVDECCQNPFLCKSYTC	133	
QY	142	VNTHGSFECYCMGDLPRNGPEPHFTTATSCITDTCGTPEVDPGVIIIGNYTSLSGS-	200	
Db	134	VNTLGSYTCQLPGF--KFIPEDPKVCTDVNCT--GQNPCHSSTHLANN-----VGSY	184	
QY	201	QVRYACRGFSV-----PDTYSSCTGLGWESPKLHCQEN-CGNPPMRHAILVGN	253	
Db	185	QCR--CRPGWQIPGSPGPNNTV--CEDVDECSGQHCQDSSTVCFN-----TVGS	232	
QY	254	HSSRLGGVARYVCQGFES-----PGKITSVCTEK--GTWRB-----SLTCTEILT	303	









```
FT DISULFID 1204 1218 BY SIMILARITY.
FT DISULFID 1220 1233 BY SIMILARITY.
FT DISULFID 1239 1250 BY SIMILARITY.
FT DISULFID 1246 1259 BY SIMILARITY.
FT DISULFID 1261 1274 BY SIMILARITY.
FT DISULFID 1280 1292 BY SIMILARITY.
FT DISULFID 1301 1316 BY SIMILARITY.
FT DISULFID 1322 1334 BY SIMILARITY.
FT DISULFID 1329 1343 BY SIMILARITY.
FT DISULFID 1345 1358 BY SIMILARITY.
FT DISULFID 1364 1377 BY SIMILARITY.
FT DISULFID 1371 1386 BY SIMILARITY.
FT DISULFID 1388 1399 BY SIMILARITY.
FT DISULFID 1405 1418 BY SIMILARITY.
FT DISULFID 1412 1427 BY SIMILARITY.
FT DISULFID 1429 1440 BY SIMILARITY.
FT DISULFID 1446 1458 BY SIMILARITY.
FT DISULFID 1453 1467 BY SIMILARITY.
FT DISULFID 1469 1482 BY SIMILARITY.
FT DISULFID 1488 1499 BY SIMILARITY.
FT DISULFID 1494 1508 BY SIMILARITY.
FT DISULFID 1510 1523 BY SIMILARITY.
FT DISULFID 1529 1540 BY SIMILARITY.
FT DISULFID 1535 1549 BY SIMILARITY.
FT DISULFID 1551 1564 BY SIMILARITY.
FT DISULFID 1647 1659 BY SIMILARITY.
FT DISULFID 1654 1668 BY SIMILARITY.
FT DISULFID 1670 1683 BY SIMILARITY.
FT DISULFID 1689 1701 BY SIMILARITY.
FT DISULFID 1696 1710 BY SIMILARITY.

Query Match 7.0%; Score 281; DB 1; Length 2907;
Best Local Similarity 28.8%; Pred. No. 8.7e-11;
Matches 83; Conservative 39; Mismatches 114; Indels 52; Gaps 15;

QY 25 RGAGAGPGLDVCA-TCHHATCOQREKIKICNVGFNGRTQCVKNEQCGATL 83
Db 1353 KGTTGCTDVDECEITGAHNCMDHSLNVPGEKSCREGWVNG-IKCIDLDEANG-TH 1410
QY 84 VCGNHTSCHNTPGGFYCICLEGYATNNKTFIPNDGTFCTDIDEC-EVSGLCRHGRCV 142
Db 1411 QCSINAQCVNTPGSRVCAKCSGF-----TGDGFTCSVDVDECAENTNLCE- 1459
QY 143 NTHGSFECYCMGVLPRNGPFPHPTDTSCTEID-CGTPPEVPDGIYNTYSSLSGQ 201
Db 1460 NVPGAYRCECMG-----FTPASDSRSQDIDECSP-----QNICVFGTCNNLPGM 1505
QY 202 VRYACREGFFSVPEDTVSSCTGLGTWESPKNLC-----QELNCGNPPMR-HAI 249
Db 1506 FHCICDDGY--ELDRTGNCCTDIDECADP-IICVNGLCVNTPGRYECNC--PPDFQLNAT 1560
QY 250 LVGNHSSRLGGVARYVCOGEPESPGKITSV-TKGTWRBETLTCTEI 297
Db 1561 GVGCVDNVRGN-----CYLKEGPRGDSLSCTEAGVGVSRSSCCSL 1603

RESULT 6
FBN2 HUMAN
ID FBN2 HUMAN STANDARD; PRT; 2911 AA.
AC P35556;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fibrillin 2 precursor.
GN FBN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94165150; PubMed=8120105;
```

```
RA Zhang H., Apfelroth S.D., Hu W., Davis E.C., Sanguineti C.,
RA Bonadio J., Mecham R.P., Ramirez F.,
RA "Structure and expression of fibrillin-2, a novel microfibrillar
RT component preferentially located in elastic matrices.";
RL J. Cell Biol. 124:855-863(1994).
RN [2]
RP SEQUENCE OF 752-1505 FROM N.A.
RX MEDLINE=91304567; PubMed=1852206;
RA Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,
RA Tsipouras P., Ramirez F., Hollister D.,
RA "Linkage of Marfan syndrome and a phenotypically related disorder to
RT two different fibrillin genes.";
RL Nature 352:330-334(1991).
RN [3]
RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.
RX MEDLINE=96083599; PubMed=7493032;
RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.,
RA "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,
RT congenital contractural arachnodactyly.";
RL Nat. Genet. 11:456-458(1995).
RN [4]
RP VARIANTS CCA HIS-1114.
RX MEDLINE=98407789; PubMed=9737771;
RA Babcock D., Gasser C., Francke U., Maalen C.,
RA "A single mutation that results in an asp-to-his substitution and
RT partial exon skipping in a family with congenital contractural
RL arachnodactyly.";
RL Hum. Genet. 103:22-28(1998).
RN [5]
RP VARIANTS CCA PHE-1141 AND TRP-1252.
RX MEDLINE=20259236; PubMed=10797416;
RA Belleh S., Zhou G., Wang M., Der Kaloustian V.M., Pagon R.A.,
RA Godfrey M.,
RA "Two novel fibrillin-2 mutations in congenital contractural
RT arachnodactyly.";
RL Am. J. Med. Genet. 92:7-12(2000).
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE
CC THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.
CC -1- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL
CC ARACHNOACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS
CC PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT EFFECT THE
CC AORTA AND THE EYES.
CC -1- SIMILARITY: Contains 47 EGF-like domains.
CC -1- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBP) domains.
CC -1- DATABASE: NAME=Elastic Fiber Homepage; NOTE=Fibrillin 2 page;
CC WWW="http://ef.wustl.edu/genes/FBN2.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U03272; AAA18950.1; -
CC EMBL; X62009; -; NOT_ANNOTATED_CDS.
CC PIR; A54105; A54105.
CC HSSP; P35555; 1EMN.
CC Genew; KGC:3604; FBN2.
CC MIM; 121050; -.
CC GO; GO:0005578; C:extracellular matrix; TAS.
CC GO; GO:0005201; F:extracellular matrix structural constituent; TAS.
CC GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
CC GO; GO:0007397; P:histogenesis and organogenesis; TAS.
CC InterPro; IPR000152; Asx hydroxyl.
CC InterPro; IPR001881; EGF-Ca.
CC InterPro; IPR001438; EGF-II.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR002212; Fibril-assoc.
CC Pfam; PF00008; EGF; 45.
CC Pfam; PF00683; TB; 9.
```

DR PRINTS; PRO0010; EGFBLD.

DR SMART; SM00179; EGF CA: 43.

DR PROSITE; PS00010; ASX HYDROXYL; 43.

DR PROSITE; PS00022; EGF\_1; 2.

DR PROSITE; PS01186; EGF\_2; 37.

DR PROSITE; PS01187; EGF CA: 42.

KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;

Repeat; Signal; Multigene family; Disease mutation; Polymorphism.

FT SIGNAL 1 28

FT CHAIN 29 2911

FT DOMAIN 111 142

FT DOMAIN 145 176

FT DOMAIN 176 207

FT DOMAIN 275 316

FT DOMAIN 317 358

FT REPEAT 359 425

FT REPEAT 493 533

FT DOMAIN 534 573

FT DOMAIN 574 615

FT DOMAIN 616 656

FT DOMAIN 657 697

FT REPEAT 698 766

FT REPEAT 767 808

FT DOMAIN 809 850

FT DOMAIN 851 890

FT DOMAIN 954 995

FT REPEAT 996 1071

FT DOMAIN 1072 1113

FT DOMAIN 1114 1156

FT DOMAIN 1157 1198

FT DOMAIN 1199 1240

FT DOMAIN 1241 1281

FT DOMAIN 1282 1323

FT DOMAIN 1324 1365

FT DOMAIN 1366 1406

FT DOMAIN 1407 1447

FT DOMAIN 1448 1489

FT DOMAIN 1490 1530

FT DOMAIN 1531 1571

FT REPEAT 1572 1648

FT REPEAT 1649 1690

FT DOMAIN 1691 1732

FT REPEAT 1733 1806

FT DOMAIN 1807 1848

FT DOMAIN 1849 1890

FT DOMAIN 1891 1932

FT DOMAIN 1933 1974

FT DOMAIN 1975 2014

FT DOMAIN 2015 2054

FT DOMAIN 2055 2096

FT REPEAT 2097 2169

FT DOMAIN 2170 2211

FT DOMAIN 2212 2251

FT DOMAIN 2252 2292

FT DOMAIN 2293 2336

FT DOMAIN 2337 2378

FT REPEAT 2379 2447

FT DOMAIN 2448 2489

FT DOMAIN 2490 2530

FT DOMAIN 2531 2569

FT DOMAIN 2570 2612

FT DOMAIN 2613 2652

FT DOMAIN 2653 2693

FT DOMAIN 2694 2733

FT DISULFID 115 124

FT DISULFID 119 130

FT DISULFID 132 141

FT DISULFID 149 159

FT DISULFID 163 164

FT DISULFID 166 175

FT DISULFID 180 190

FT DISULFID 194 195

FT DISULFID 197 206

FT DISULFID 279 291

FT DISULFID 286 300

FT DISULFID 302 315

FT DISULFID 321 333

FT DISULFID 328 342

FT DISULFID 344 357

FT DISULFID 497 509

FT DISULFID 504 518

FT DISULFID 520 532

FT DISULFID 538 548

FT DISULFID 543 557

FT DISULFID 559 572

FT DISULFID 578 590

FT DISULFID 585 599

FT DISULFID 601 614

FT DISULFID 620 631

FT DISULFID 626 640

FT DISULFID 642 655

FT DISULFID 661 672

FT DISULFID 667 681

FT DISULFID 683 696

FT DISULFID 771 783

FT DISULFID 778 792

FT DISULFID 794 807

FT DISULFID 813 825

FT DISULFID 820 834

FT DISULFID 836 849

FT DISULFID 855 865

FT DISULFID 860 874

FT DISULFID 876 889

FT DISULFID 958 970

FT DISULFID 965 979

FT DISULFID 981 994

FT DISULFID 1076 1088

FT DISULFID 1083 1097

FT DISULFID 1099 1112

FT DISULFID 1118 1130

FT DISULFID 1125 1139

FT DISULFID 1141 1155

FT DISULFID 1161 1173

Query Match 6.9%; Score 278.5; DB 1; Length 2911;

Best Local Similarity 28.8%; Pred. No. 1.3e-10;

Matches 90; Conservative 30; Mismatches 126; Indels 67; Gaps 17;

QY 2 GRGWADGSPRRLLPLLLLLGLARG-----RAGAPGP--DGLDV---CATCHEHAT-- 47

Db 449 GYGFPGTG----FIPPGNGFSPGVGGAGVAGGQGPITGLTILNQTIDICKHANLC 504

QY 48 ---CQOREGKICICNYGVGVNGRTQCVDKNECFGATLVCGNHTSCHNTFGFYICL 103

Db 505 LNGRCIPTVSSYRCECMGKQDANGDCIDVDEC---TSNPCTN-GDCVNTFGSYCKCH 560

QY 104 EGYRATNNKTFIPNDGTCTDIDECEVSLGRHGRCVNTHGSEPCYCMQGYLPRNGPE 163

Db 561 AGFQRTPTKQA-----CIDIDECIQNGVLCKNGRCVNSDGSFQCICNAG----- 604

QY 164 PFHPTTATSCTEID-CGTPPEVPDGYIIGNYTSLSGQSVRYACREGFFSVPEDTVSSCT 222

Db 605 -FELTDDKNCVDHDECTTNNCLMGWCINE-----DGSFKCICKPGFVLAPNGRY--CT 656

QY 223 GLGTWESPXL-----HC-----QELINGNPP-----EMRHALLVGNH--SSRLGQGVARYVCOE 268

Db 657 DVDECOPTGFCMNGHCINSESGFRCDPPGLAVGMDGRVCVDTHMRSTCYGIGKGVCR 716

QY 269 GFESPGGKITSV 281

Db 717 PF--PGAVTKSEC 727

RESULT 7

ID UROM HUMAN STANDARD; PRT; 640 AA.

AC P07911;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Uromodulin precursor (Tamm-Horsfall urinary glycoprotein) (THP).  
GN UMOD.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxid=9606;  
RN [1]  
RP MEDLINE=87177970; PubMed=3453112;  
RA Pennica D., Kohr W.-J., Kuang W.-J., Glaister D., Aggarwal B.B.,  
Chen E.Y., Goeddel D.V.;  
RT "Identification of human uromodulin as the Tamm-Horsfall urinary  
glycoprotein";  
RL Science 236:83-88(1987).  
RN [2]  
RP MEDLINE=87319675; PubMed=3498215;  
RA Hession C., Decker J.M., Sherblom A.P., Kumar S., Yue C.C.,  
Mataliano R.J., Tizard R., Kawashima E., Schmeissner U.,  
Helletky S., Chow E.P., Burne C.A., Shaw A., Muchmore A.V.;  
RT "Uromodulin (Tamm-Horsfall glycoprotein): a renal ligand for  
lymphokines";  
RL Science 237:1479-1484(1987).  
RN [3]  
RP GPI-ANCHOR.  
RA MEDLINE=91065873; PubMed=2249987;  
RX Rindler M.J., Naik S.S., Li N., Hoops T.C., Peraldi M.-N.;  
RT "Uromodulin (Tamm-Horsfall glycoprotein/aromucoid) is a  
phosphatidylinositol-linked membrane protein";  
RJ J. Biol. Chem. 265:20784-20789(1990).  
CC -!- FUNCTION: NOT KNOWN. MAY PLAY A ROLE IN REGULATING THE CIRCULATING  
ACTIVITY OF CYTOKINES AS IT BINDS TO IL-1, IL-2 AND TNF WITH HIGH  
AFFINITY.  
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR,  
THEN CLEAVED TO PRODUCE A SOLUBLE FORM WHICH IS SECRETED IN  
URINE.  
CC -!- TISSUE SPECIFICITY: SYNTHESIZED BY THE KIDNEYS AND IS THE MOST  
ABUNDANT PROTEIN IN NORMAL HUMAN URINE.  
CC -!- SIMILARITY: Contains 3 EGF-like domains.  
CC -!- SIMILARITY: Contains 1 ZP domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M15881; AAA36798.1; -;  
DR EMBL; M17778; AAA36799.1; -;  
DR PIR; A30452; A30452.  
DR HSP; P07204; IADK.  
DR GlycoSuiteDB; P07911; -;  
DR Genew; HGNC:12559; UMOD.  
DR MIM; 191845; -;  
DR GO; GO:0015025; F:GPI-anchored membrane-bound receptor; TAS.  
DR GO; GO:0006968; P:cellular defense response; TAS.  
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
DR InterPro; IPR000152; Asx hydroxyl.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF\_Like.  
DR InterPro; IPR001507; Endoglin/CD105.  
DR Pfam; PF00008; EGF; 3.  
DR Pfam; PF00100; zona pellucida; 1.  
DR PRINTS; PR00023; ZPELLUCIDA.  
DR SMART; SM00179; EGF\_CA; 2.  
DR SMART; SM00241; ZP\_1.  
DR PROSITE; PS00682; ZP\_DOMAIN; 1.

DR PROSITE; PS00010; ASX HYDROXYL; 2.  
DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
DR PROSITE; PS01186; EGF\_2; 3.  
DR PROSITE; PS01187; EGF\_CA; 2.  
KW Glycoprotein; Signal; Membrane; GPI-anchor; EGF-like domain.  
FT SIGNAL 1 24  
FT CHAIN 25 640 UROMODULIN.  
FT DOMAIN 28 64 EGF-LIKE 1.  
FT DOMAIN 65 107 EGF-LIKE 2.  
FT DOMAIN 108 149 EGF-LIKE 3.  
FT DOMAIN 334 585 ZP.  
FT DISULFID 32 41 BY SIMILARITY.  
FT DISULFID 35 50 BY SIMILARITY.  
FT DISULFID 52 63 BY SIMILARITY.  
FT DISULFID 69 83 BY SIMILARITY.  
FT DISULFID 77 92 BY SIMILARITY.  
FT DISULFID 94 106 BY SIMILARITY.  
FT DISULFID 112 126 BY SIMILARITY.  
FT DISULFID 120 135 BY SIMILARITY.  
FT DISULFID 137 148 BY SIMILARITY.  
FT CARBOHYD 38 38 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 76 76 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 80 80 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .).  
FT CARBOHYD 275 275 N-LINKED (GLCNAC. .).  
FT CARBOHYD 322 322 /ETID-CAR 000178.  
FT CARBOHYD 322 322 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 396 396 N-LINKED (GLCNAC. .).  
FT CONFLICT 565 565 H -> D (IN REF. 2).  
SQ SEQUENCE 640 AA; 69760 MW; D26A07A76353AE48 CRC64;  
Query Match 6.9%; Score 277.5; DB 1; Length 640;  
Best Local Similarity 23.4%; Pred. No. 2-2e-11;  
Matches 163; Conservative 77; Mismatches 217; Indels 239; Gaps 38;  
QY 39 CATCEHATCOQREGKKICICNYGVFVNGRTQCKNECQFATLVCGNHTSCHTTPGQF 98  
DB 32 CSECHSNAICTDEAVTTCQEGFTGDGLT-CVDLDECAIFCAHNSANSCVNPGBF 90  
QY 99 YCICLGXYATNNKTFIPNDGTFCTDIDECVSLG--CRHGRCVNTHGSPCYCMQGY 156  
DB 91 SCVCPGFRU-----SPGLGCTDVECAEPGLSHCHALATCVNVVGSYLCVCPAGY 141  
QY 157 -----LPRN-----GPEPHPTTD----- 170  
DB 142 RGDGHECECSGPGGLDCVPEGDALVCADPCQAHRTLDYWRSTEGYACDITDLRG 201  
QY 171 -----ATSCTEI-DC-----GTPPEVPDGYIIGNYTS-----SLG 199  
DB 202 WYFVQGGGARMAETCPVLRNNTAAPMWLNGTHPSDEGIVSRKACAHWSGHCCLDAS 261  
QY 200 SQVRYACREGFF-----SVPEDTVSSCTG-----LGTWE-----SPKLHC--QE 236  
DB 262 VQVK-ACAGYVYVNLTAPECHLAYCTDPSSVEGTCECSIDEDCKSNNGRWHCCQCKD 320  
QY 237 INCNPPEMRHAILVGNHSLGGVARYVQ---EGFSPGGKI-----TSVCT-----E 283  
DB 321 FNITDISLLEHLECCANDMK---VSLGKCOLKSLGFD---KVFYLSDRSCGNDRD 373  
QY 284 KGTWRESTLT-----CTEILTKINDVSLFNDDTC-VRWQINSRRINPKISYV----- 328  
DB 374 NRDM-VSVVTPARDGPGGVLTENETHATYNTLYLADBIIRDLNLIKINFACSYPLDMK 432  
QY 329 ISIKGQRLDPMES---VRSETVNLITDSTFPVCLALYFGTYNTVNIATPPRSMPAVI 385  
DB 433 VSLK-TALQPMVSALNIRVGTGMFT-----VRMALFOTPSYV-----QP--- 471  
QY 386 GFOTAEVDLLEDGGSFNISIFNETCKLNRRSRKVGSEHMYQFTVLGQRYLANFHSATS 445  
DB 472 -YQSSVTL-----STEAFLYVGTMLDGGDLR-----FALLMTNCAFPSSNATD 516  
QY 446 --FNFTTREQVPVCLDLPTTDTYVNVTL--LRSPKRHSVQITATPPAVKQVITNISG 501

Db 517 PLKFIIDRCF-----HTRDSTQVWNGSSQGRFVQM----- 552  
QY 502 FNETCLRWRSIKTADME-----EMYLEHFWG:RWYQKEFAQEMTNISSSSRDPEVCIDLR 557  
Db 553 -----PREFAGNYDLVYLHCEVILCDTMAE:KC--KPTCSGTRPR-SGSVIDQSRVLNLG 602  
QY 558 POTNNV-----SURLSELPPVVIS:TTQIT 584  
Db 603 PITRGVQVATVSRAFFSSGLLKWLPPLLS:TLTLT 638  
RESULT 8  
EMR1\_HUMAN  
ID EMR1\_HUMAN STANDARD; PRT; 88; AA.  
AC Q14246;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cell surface glycoprotein EMR1 precursor (EMR1 hormone receptor).  
GN EMR1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.  
OX NCBI Taxid=9606;  
RN [1]  
RP MEDLINE=95324926; PubMed=7601460;  
RA Baud V., Chisoe S.L., Viegas-Pequignot E., Diriong S., N'Guyen V.C.,  
RA "EMR1, an unusual member in the family of hormone receptors with  
RT seven transmembrane segments.";  
RL Genomics 26:334-344 (1995).  
CC -!- FUNCTION: PROBABLY INVOLVED IN CELLULAR RESPONSE TO A HORMONE OR  
CC -!- AN INTERACTION WITH A PROTEIN LIGAND.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- TISSUE SPECIFICITY: WIDE EXPRESSION; INCREASED LEVELS IN  
CC PERIPHERAL BLOOD MONONUCLEAR CELLS.  
CC -!- PTM: N- AND O-GLYCOSYLATED; (POSSIBLE).  
CC -!- SIMILARITY: Contains 6 EGF-like domains.  
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF 3-PROTEIN COUPLED RECEPTORS.  
CC -!- SIMILARITY: Contains 1 GPS domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X81479; CAA57232.1; -  
CC PIR; A57172; A57172.  
CC HSP; P00736; IAPQ.  
CC Genew; HGNC:3336; EMR1.  
CC MIM; 600493; -  
CC GO; GO:0005887; C: integral to plasma membrane; TAS.  
CC GO; GO:0004930; F: G-protein coupled receptor activity; TAS.  
CC GO; GO:0007155; P: cell adhesion; TAS.  
CC InterPro; IPR000152; Asx hydroxyl.  
CC InterPro; IPR001881; EGF\_Ca.  
CC InterPro; IPR006209; EGF\_like.  
CC InterPro; IPR000832; GPCR\_secretin.  
CC Pfam; PF00002; 7tm\_2; 1.  
CC Pfam; PF00008; EGF; 5.  
CC Pfam; PF01825; GPS; 1.  
CC PRINTS; PR00249; GPCRSSECRETIN.  
CC SMART; SM00179; EGF\_CA; 5.  
CC SMART; SM00303; GPS\_1.  
CC PROSITE; PS00010; ASX\_HYDROXYL; 6.  
CC PROSITE; PS01186; EGF\_2; 2.  
CC PROSITE; PS01187; EGF\_CA; 5.  
CC PROSITE; PS50221; GPS; 1.

DR PROSITE; PS00650; G-PROTEIN RECEPTOR P2\_2; 1.  
DR PROSITE; PS0261; G-PROTEIN RECEPTOR P2\_4; 1.  
KW G-protein coupled receptor; Transmembrane; Receptor; Glycoprotein;  
KW EGF-like domain; Repeat; Signal.  
FT SIGNAL 1  
FT CHAIN 18  
FT DOMAIN 18 886 CELL SURFACE GLYCOPROTEIN EMR1.  
FT DOMAIN 18 599 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 600 627 POTENTIAL.  
FT DOMAIN 628 634 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 635 656 POTENTIAL.  
FT DOMAIN 657 666 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 667 690 POTENTIAL.  
FT DOMAIN 691 709 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 710 731 POTENTIAL.  
FT DOMAIN 732 747 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 748 776 POTENTIAL.  
FT DOMAIN 777 794 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 795 814 POTENTIAL.  
FT DOMAIN 815 829 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 830 852 POTENTIAL.  
FT DOMAIN 853 886 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 31 79 EGF-LIKE 1.  
FT DOMAIN 80 131 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 132 171 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 172 213 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 221 267 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 268 316 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 317 599 GPS.  
FT DOMAIN 317 599 SER/THR-RICH.  
FT DISULFID 35 47 BY SIMILARITY.  
FT DISULFID 41 56 BY SIMILARITY.  
FT DISULFID 58 78 BY SIMILARITY.  
FT DISULFID 84 97 BY SIMILARITY.  
FT DISULFID 91 106 BY SIMILARITY.  
FT DISULFID 108 130 BY SIMILARITY.  
FT DISULFID 136 148 BY SIMILARITY.  
FT DISULFID 142 157 BY SIMILARITY.  
FT DISULFID 159 170 BY SIMILARITY.  
FT DISULFID 176 188 BY SIMILARITY.  
FT DISULFID 182 197 BY SIMILARITY.  
FT DISULFID 199 212 BY SIMILARITY.  
FT DISULFID 225 235 BY SIMILARITY.  
FT DISULFID 229 244 BY SIMILARITY.  
FT DISULFID 246 266 BY SIMILARITY.  
FT DISULFID 272 285 BY SIMILARITY.  
FT DISULFID 279 294 BY SIMILARITY.  
FT DISULFID 296 315 BY SIMILARITY.  
FT CARBOHYD 94 94 N-LINKED (GLCNAC... (POTENTIAL).  
FT CARBOHYD 99 99 N-LINKED (GLCNAC... (POTENTIAL).  
FT CARBOHYD 127 127 N-LINKED (GLCNAC... (POTENTIAL).  
FT CARBOHYD 167 167 N-LINKED (GLCNAC... (POTENTIAL).  
FT CARBOHYD 189 189 N-LINKED (GLCNAC... (POTENTIAL).  
FT CARBOHYD 194 194 N-LINKED (GLCNAC... (POTENTIAL).  
FT CARBOHYD 232 232 N-LINKED (GLCNAC... (POTENTIAL).  
FT CARBOHYD 258 258 N-LINKED (GLCNAC... (POTENTIAL).  
FT CARBOHYD 312 312 N-LINKED (GLCNAC... (POTENTIAL).  
FT CARBOHYD 366 366 N-LINKED (GLCNAC... (POTENTIAL).  
FT CARBOHYD 375 375 N-LINKED (GLCNAC... (POTENTIAL).  
FT CARBOHYD 448 448 N-LINKED (GLCNAC... (POTENTIAL).  
FT CARBOHYD 661 661 N-LINKED (GLCNAC... (POTENTIAL).  
SQ SEQUENCE 886 AA; 97680 MW; 7456CA56F8624D99 CRC64;

Query Match 6.8%; Score 274.5; DB 1; Length 886;

Best Local Similarity 22.8%; Pred. No. 5.3e-11;

Matches 89; Conservative 53; Mismatches 125; Indels 123; Gaps 17;

QY 42 CHEHATCOQREGKKKICICNNGFV-GNGR-----TQVDKNECQFATLVCGNHTSCHN 93

Db 41 CPAYATCTNVDYSYCTCKQGLSSNGQNFHFKDPGVRCXIDECSS-QSPQCPGNSSCKN 99

QY 94 TPGGFYICILEGYRANNKTFIPNDGTF-CITDICEVSGLCRRGRCVHTHGFECVC 152

Db 100 LSGRYKSCLDGFSPTGNDWVPGKGNFSCDINECLTSRVCPHSDCVMSGYSKSC 159

Qy 153 MDGLVPRNGPEPHPTDTSCTEI-DCGTEPEVPD-----GYLIGNYTSLSGQVYACR 207  
Db 160 QVGFISRN-----STCEDVNECAPACPEHATCNNTVGNYS-FCN 200  
Qy 208 EGFPSVPEDTVSSCTGLGTWESPKLHCQEI-----CGNPPEMRHAILVGNHSS 256  
Db 201 PGES---SSGHLSCQL-----KASCEDIDECTEMPCINSTCTNP-----239  
Qy 257 RLGVARYVCOGFESPGGKI-----TSVCTE-----283  
Db 240 ---GSYFCTCHPGFAPSPGQLNFETDQGVCEKIDIDECRQDPSTCGPNSICTNALGSYSCG 296  
Qy 284 -----KGWRESTITCTEITLKINDVSLENDTCVRWQINRRINP-----KISV 328  
Db 297 IVGHPHPESQKQDNFSQORVLFKCKEDVFNKQIQOQCGTAVKPAYVSFCAQINNI 356  
Qy 329 ISIKQBLDPMESVREETVAL--TTDSRTP 356  
Db 357 FSV---LDKVCENKTTWSLKNTTESFVP 382

## RESULT 9

FBNI\_PIG STANDARD; PRT; 2871 AA.  
AC Q9TV36;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Fibrillin 1 precursor.  
GN FBNI.  
OS Sus scrofa (Pig).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBT\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=99156858; PubMed=10036187;  
RA Biery N.J., Eldadah Z.A., Moore C.S., Statten G., Spencer F.,  
RA Dietz H.C.;  
RT "Revised genomic organization of FBNI and significance for regulated  
gene expression";  
RL Genomics 56:70-77(1999).  
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE  
LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.  
CC -!- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER  
FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE  
MICROFIBRILS.  
CC -!- SIMILARITY: Contains 47 EGF-like domains.  
CC -!- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBP) domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; AF073800; AAD50328.1; --  
DR HSPF; F35555; IAPF.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR002212; Fibril-assoc.  
DR Pfam; PF00008; EGF; 45.  
DR Pfam; PF00683; TB; 9.  
DR SMART; SM00179; EGF\_Ca; 40.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 41.  
DR PROSITE; PS00022; EGF\_1; 2.  
DR PROSITE; PS01186; EGF\_2; 36.

DR PROSITE; PS01187; EGF\_Ca; 41.  
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  
KW Repeat; Signal; Multigene family.  
FT SIGNAL 1 27  
FT CHAIN 28 2871  
FT FIBRILLIN 1.  
FT EGF-LIKE 1.  
FT EGF-LIKE 2.  
FT EGF-LIKE 3.  
FT EGF-LIKE 4.  
FT EGF-LIKE 5.  
FT EGF-LIKE 6.  
FT EGF-LIKE 7.  
FT EGF-LIKE 8.  
FT EGF-LIKE 9.  
FT EGF-LIKE 10.  
FT EGF-LIKE 11.  
FT EGF-LIKE 12.  
FT EGF-LIKE 13.  
FT EGF-LIKE 14.  
FT EGF-LIKE 15.  
FT EGF-LIKE 16.  
FT EGF-LIKE 17.  
FT EGF-LIKE 18.  
FT EGF-LIKE 19.  
FT EGF-LIKE 20.  
FT EGF-LIKE 21.  
FT EGF-LIKE 22.  
FT EGF-LIKE 23.  
FT EGF-LIKE 24.  
FT EGF-LIKE 25.  
FT EGF-LIKE 26.  
FT EGF-LIKE 27.  
FT EGF-LIKE 28.  
FT EGF-LIKE 29.  
FT EGF-LIKE 30.  
FT EGF-LIKE 31.  
FT EGF-LIKE 32.  
FT EGF-LIKE 33.  
FT EGF-LIKE 34.  
FT EGF-LIKE 35.  
FT EGF-LIKE 36.  
FT EGF-LIKE 37.  
FT EGF-LIKE 38.  
FT EGF-LIKE 39.  
FT EGF-LIKE 40.  
FT EGF-LIKE 41.  
FT EGF-LIKE 42.  
FT EGF-LIKE 43.  
FT EGF-LIKE 44.  
FT EGF-LIKE 45.  
FT EGF-LIKE 46.  
FT EGF-LIKE 47.  
FT EGF-LIKE 48.  
FT EGF-LIKE 49.  
FT EGF-LIKE 50.  
FT EGF-LIKE 51.  
FT EGF-LIKE 52.  
FT EGF-LIKE 53.  
FT EGF-LIKE 54.  
FT EGF-LIKE 55.  
FT EGF-LIKE 56.  
FT EGF-LIKE 57.  
FT EGF-LIKE 58.  
FT EGF-LIKE 59.  
FT EGF-LIKE 60.  
FT EGF-LIKE 61.  
FT EGF-LIKE 62.  
FT EGF-LIKE 63.  
FT EGF-LIKE 64.  
FT EGF-LIKE 65.  
FT EGF-LIKE 66.  
FT EGF-LIKE 67.  
FT EGF-LIKE 68.  
FT EGF-LIKE 69.  
FT EGF-LIKE 70.  
FT EGF-LIKE 71.  
FT EGF-LIKE 72.  
FT EGF-LIKE 73.  
FT EGF-LIKE 74.  
FT EGF-LIKE 75.  
FT EGF-LIKE 76.  
FT EGF-LIKE 77.  
FT EGF-LIKE 78.  
FT EGF-LIKE 79.  
FT EGF-LIKE 80.  
FT EGF-LIKE 81.  
FT EGF-LIKE 82.  
FT EGF-LIKE 83.  
FT EGF-LIKE 84.  
FT EGF-LIKE 85.  
FT EGF-LIKE 86.  
FT EGF-LIKE 87.  
FT EGF-LIKE 88.  
FT EGF-LIKE 89.  
FT EGF-LIKE 90.  
FT EGF-LIKE 91.  
FT EGF-LIKE 92.  
FT EGF-LIKE 93.  
FT EGF-LIKE 94.  
FT EGF-LIKE 95.  
FT EGF-LIKE 96.  
FT EGF-LIKE 97.  
FT EGF-LIKE 98.  
FT EGF-LIKE 99.  
FT EGF-LIKE 100.  
FT EGF-LIKE 101.  
FT EGF-LIKE 102.  
FT EGF-LIKE 103.  
FT EGF-LIKE 104.  
FT EGF-LIKE 105.  
FT EGF-LIKE 106.  
FT EGF-LIKE 107.  
FT EGF-LIKE 108.  
FT EGF-LIKE 109.  
FT EGF-LIKE 110.  
FT EGF-LIKE 111.  
FT EGF-LIKE 112.  
FT EGF-LIKE 113.  
FT EGF-LIKE 114.  
FT EGF-LIKE 115.  
FT EGF-LIKE 116.  
FT EGF-LIKE 117.  
FT EGF-LIKE 118.  
FT EGF-LIKE 119.  
FT EGF-LIKE 120.  
FT EGF-LIKE 121.  
FT EGF-LIKE 122.  
FT EGF-LIKE 123.  
FT EGF-LIKE 124.  
FT EGF-LIKE 125.  
FT EGF-LIKE 126.  
FT EGF-LIKE 127.  
FT EGF-LIKE 128.  
FT EGF-LIKE 129.  
FT EGF-LIKE 130.  
FT EGF-LIKE 131.  
FT EGF-LIKE 132.  
FT EGF-LIKE 133.  
FT EGF-LIKE 134.  
FT EGF-LIKE 135.  
FT EGF-LIKE 136.  
FT EGF-LIKE 137.  
FT EGF-LIKE 138.  
FT EGF-LIKE 139.  
FT EGF-LIKE 140.  
FT EGF-LIKE 141.  
FT EGF-LIKE 142.  
FT EGF-LIKE 143.  
FT EGF-LIKE 144.  
FT EGF-LIKE 145.  
FT EGF-LIKE 146.  
FT EGF-LIKE 147.  
FT EGF-LIKE 148.  
FT EGF-LIKE 149.  
FT EGF-LIKE 150.  
FT EGF-LIKE 151.  
FT EGF-LIKE 152.  
FT EGF-LIKE 153.  
FT EGF-LIKE 154.  
FT EGF-LIKE 155.  
FT EGF-LIKE 156.  
FT EGF-LIKE 157.  
FT EGF-LIKE 158.  
FT EGF-LIKE 159.  
FT EGF-LIKE 160.  
FT EGF-LIKE 161.  
FT EGF-LIKE 162.  
FT EGF-LIKE 163.  
FT EGF-LIKE 164.  
FT EGF-LIKE 165.  
FT EGF-LIKE 166.  
FT EGF-LIKE 167.  
FT EGF-LIKE 168.  
FT EGF-LIKE 169.  
FT EGF-LIKE 170.  
FT EGF-LIKE 171.  
FT EGF-LIKE 172.  
FT EGF-LIKE 173.  
FT EGF-LIKE 174.  
FT EGF-LIKE 175.  
FT EGF-LIKE 176.  
FT EGF-LIKE 177.  
FT EGF-LIKE 178.  
FT EGF-LIKE 179.  
FT EGF-LIKE 180.  
FT EGF-LIKE 181.  
FT EGF-LIKE 182.  
FT EGF-LIKE 183.  
FT EGF-LIKE 184.  
FT EGF-LIKE 185.  
FT EGF-LIKE 186.  
FT EGF-LIKE 187.  
FT EGF-LIKE 188.  
FT EGF-LIKE 189.  
FT EGF-LIKE 190.  
FT EGF-LIKE 191.  
FT EGF-LIKE 192.  
FT EGF-LIKE 193.  
FT EGF-LIKE 194.  
FT EGF-LIKE 195.  
FT EGF-LIKE 196.  
FT EGF-LIKE 197.  
FT EGF-LIKE 198.  
FT EGF-LIKE 199.  
FT EGF-LIKE 200.  
FT EGF-LIKE 201.  
FT EGF-LIKE 202.  
FT EGF-LIKE 203.  
FT EGF-LIKE 204.  
FT EGF-LIKE 205.  
FT EGF-LIKE 206.  
FT EGF-LIKE 207.  
FT EGF-LIKE 208.  
FT EGF-LIKE 209.  
FT EGF-LIKE 210.  
FT EGF-LIKE 211.  
FT EGF-LIKE 212.  
FT EGF-LIKE 213.  
FT EGF-LIKE 214.  
FT EGF-LIKE 215.  
FT EGF-LIKE 216.  
FT EGF-LIKE 217.  
FT EGF-LIKE 218.  
FT EGF-LIKE 219.  
FT EGF-LIKE 220.  
FT EGF-LIKE 221.  
FT EGF-LIKE 222.  
FT EGF-LIKE 223.  
FT EGF-LIKE 224.  
FT EGF-LIKE 225.  
FT EGF-LIKE 226.  
FT EGF-LIKE 227.  
FT EGF-LIKE 228.  
FT EGF-LIKE 229.  
FT EGF-LIKE 230.  
FT EGF-LIKE 231.  
FT EGF-LIKE 232.  
FT EGF-LIKE 233.  
FT EGF-LIKE 234.  
FT EGF-LIKE 235.  
FT EGF-LIKE 236.  
FT EGF-LIKE 237.  
FT EGF-LIKE 238.  
FT EGF-LIKE 239.  
FT EGF-LIKE 240.  
FT EGF-LIKE 241.  
FT EGF-LIKE 242.  
FT EGF-LIKE 243.  
FT EGF-LIKE 244.  
FT EGF-LIKE 245.  
FT EGF-LIKE 246.  
FT EGF-LIKE 247.  
FT EGF-LIKE 248.  
FT EGF-LIKE 249.  
FT EGF-LIKE 250.  
FT EGF-LIKE 251.  
FT EGF-LIKE 252.  
FT EGF-LIKE 253.  
FT EGF-LIKE 254.  
FT EGF-LIKE 255.  
FT EGF-LIKE 256.  
FT EGF-LIKE 257.  
FT EGF-LIKE 258.  
FT EGF-LIKE 259.  
FT EGF-LIKE 260.  
FT EGF-LIKE 261.  
FT EGF-LIKE 262.  
FT EGF-LIKE 263.  
FT EGF-LIKE 264.  
FT EGF-LIKE 265.  
FT EGF-LIKE 266.  
FT EGF-LIKE 267.  
FT EGF-LIKE 268.  
FT EGF-LIKE 269.  
FT EGF-LIKE 270.  
FT EGF-LIKE 271.  
FT EGF-LIKE 272.  
FT EGF-LIKE 273.  
FT EGF-LIKE 274.  
FT EGF-LIKE 275.  
FT EGF-LIKE 276.  
FT EGF-LIKE 277.  
FT EGF-LIKE 278.  
FT EGF-LIKE 279.  
FT EGF-LIKE 280.  
FT EGF-LIKE 281.  
FT EGF-LIKE 282.  
FT EGF-LIKE 283.  
FT EGF-LIKE 284.  
FT EGF-LIKE 285.  
FT EGF-LIKE 286.  
FT EGF-LIKE 287.  
FT EGF-LIKE 288.  
FT EGF-LIKE 289.  
FT EGF-LIKE 290.  
FT EGF-LIKE 291.  
FT EGF-LIKE 292.  
FT EGF-LIKE 293.  
FT EGF-LIKE 294.  
FT EGF-LIKE 295.  
FT EGF-LIKE 296.  
FT EGF-LIKE 297.  
FT EGF-LIKE 298.  
FT EGF-LIKE 299.  
FT EGF-LIKE 300.











FT	DOMAIN	592	641	GPS.	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	506	508	BY SIMILARITY.	
FT	DISULFID	36	48	BY SIMILARITY.	
FT	DISULFID	42	57	BY SIMILARITY.	
FT	DISULFID	59	79	BY SIMILARITY.	
FT	DISULFID	85	98	BY SIMILARITY.	
FT	DISULFID	92	107	BY SIMILARITY.	
FT	DISULFID	109	131	BY SIMILARITY.	
FT	DISULFID	137	149	BY SIMILARITY.	
FT	DISULFID	143	158	BY SIMILARITY.	
FT	DISULFID	160	171	BY SIMILARITY.	
FT	DISULFID	177	189	BY SIMILARITY.	
FT	DISULFID	183	198	BY SIMILARITY.	
FT	DISULFID	200	220	BY SIMILARITY.	
FT	DISULFID	226	239	BY SIMILARITY.	
FT	DISULFID	233	248	BY SIMILARITY.	
FT	DISULFID	250	270	BY SIMILARITY.	
FT	DISULFID	276	286	BY SIMILARITY.	
FT	DISULFID	280	295	BY SIMILARITY.	
FT	DISULFID	297	317	BY SIMILARITY.	
FT	DISULFID	323	336	BY SIMILARITY.	
FT	DISULFID	330	345	BY SIMILARITY.	
FT	DISULFID	347	366	BY SIMILARITY.	
FT	CARBOHYD	148	148	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	167	167	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	229	229	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	269	269	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	283	283	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	405	405	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	417	417	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	474	474	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	498	498	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	706	706	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	SEQUENCE	931 AA;	102129 MW;	52963A667E8B76B5 CRC64;	

Query Match

Best Local Similarity 23.9%; Pred. No. 5.4e-10;

Matches 123; Conservative 54; Mismatches 192; Indels 146; Gaps 27;

QY	20	LLGLA-----RGAAGAPGD-----GLDVCAT-----CHEAFCQREKKICIC	59
DB	101	ILGRKCSCLRGFSSTGKDWILGSLDNFLADVDECLFGICPKYCNCSNVGSYSCTC	160
QY	60	NYGVFGNGRTQCVKNECFGATVCGNHTSCNTPGGFYCICLEGYRATNN-----	111
DB	161	QPGFVLNG-SICEDEECV--TRDVCPEHAQCHNTLGSYYCTCNGLSSGGGPMFQGLD	217
QY	112	-----NKTFI-----PNDGTFCTDIDEC	129
DB	218	ESCEDVDECSRNTLCGPTFCINTLGSYSQCPAGFSLFTFQILGHPADGN-CTDIDEC	276
QY	130	EVSGLCRHGRCVNTGHSPECYCMGVLPRNGPPEPH-----PTDATSCTE--IDCG---	180
DB	277	D--DTCPNLSSCNTTIGSYFCTCHPGFASNGQLNFKDLVTCEDIDECTQPLQGLNS	334
QY	181	TPPEVDPGYIIG---NYTSSLSGQ---VRYACREGFPFVPEDTVSSCTGLGWESPKL-H	233
DB	335	VCTNVPGSYICGLPDPFQMDPEGSQGVGNFCKRILFKCKEDLI-----LQSEIQIQ	386
QY	234	QOINCNPNPEMRAHVLGNHSSRLGGVAP/VQCE-----GPSPGKITSVCTEKGW-	287
DB	387	COAVQGRDLGYASFCLVNAFTILDN---TCENKSAPVLSQSAATSVSLVEQRTWTF	442
QY	288	---RESTLTCTEILTATKINDVSLFNDCVWRWINSRINPKISYIVISIKGRQLDPMESVRE	344
DB	443	ELSKETSTLGTILLETVESTMLAALLIPENASQMIQTE---YLDIESKVINE-ECKEN	498
QY	345	ETVNLTTDSRTPEV-CUALYPGNTYTNISTAPRPSMPAVIGFOAEVDL-----LEDGQ	399
DB	499	ESINLAARGDMVNGCFII-----KESVSGAPG---VAFVSFAHVESVILNERFPEDGQ	549
QY	408	SENI-SIFNETCLNRRSRKVGSEHMYQFT/LGOR	434

DB	550	SPR-----KLNMNSRVVGG-----TVTGEK	569
DB	RESULT 13		
DB	LTBS MOUSE		
AC	Q8CGI8; Q8BNW7; Q8C7F5; Q8CIR0;	PRT; 1389 AA.	
DT	15-SEP-2003 (Rel. 42, Created)		
DT	15-SEP-2003 (Rel. 42, Last sequence update)		
DE	Latent transforming growth factor beta binding protein, isoform 1S precursor (LTBP-1) (Transforming growth factor beta-1 binding protein 1) (TGF-beta1-BP-1).		
GN	LTBP1..		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	PubMed=12711388;		
RA	Noguera I., Obata H., Gualandris A., Cowin P., Rifkin D.B.;		
RT	"Molecular cloning of the mouse ltbp-1 gene reveals tissue specific expression of alternatively spliced forms.";		
RL	Gene 308:31-41(2003).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	STRAIN=129/SvJ;		
RC	PubMed=12711389;		
RA	Weiskirchen R., Moser M., Guenther K., Weiskirchen S., Gressner A.M.;		
RT	"The murine latent transforming growth factor-beta binding protein (ltbp-1) is alternatively spliced, and maps to a region syntenic to human chromosome 2p21-22.";		
RL	Gene 308:43-52(2003).		
RN	[3]		
RP	SEQUENCE OF 788-1389 FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Aorta, Liver, and Vein;		
RX	PubMed=12466851;		
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,		
RA	Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,		
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schorbach C., Gojobori T.,		
RA	Bardelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,		
RA	Schrim L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,		
RA	Blake J.A., Bradt D., Brusic V., Clothia C., Corbani L.E., Cousins S.,		
RA	Dalla E., Dragani T.A., Fietcher C.F., Forrest A., Gough J.,		
RA	Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,		
RA	Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,		
RA	Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,		
RA	Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,		
RA	Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,		
RA	Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,		
RA	Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,		
RA	Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,		
RA	Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,		
RA	Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,		
RA	Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,		
RA	Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,		
RA	Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,		
RA	Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,		
RA	Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,		
RA	Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,		
RA	Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,		
RA	Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,		
RA	Birney E., Hayashizaki Y.;		
RT	"Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";		
RL	Nature 420:563-573(2002).		
CC	-!- SUBUNIT: The large latent complex of TGF-beta1 from platelets is composed of the TGF-beta1 molecule noncovalently associated with a disulfide-bonded complex of a dimer of the N-terminal propeptide of the TGF-beta1 precursor and a third component denoted TGF-beta1-BP. TGF-beta1-BP does not bind directly to active TGF-BETA1. Binds to fibrillin (By similarity).		



Db 894 FHCVEQG-----FSISADORTCEDIECVNNTVCDSH---GFCNTAGSFRCLCY 941  
QY 208 EGFSPVEDTSSCTGLTWESPXKHCQET:CGNPPMRHAILVGNHSSRLGSGVARYVCQ 267  
Db 942 QG-FQAPQD-----GQG-----CQDV-----NECELLSG-----VCG 967  
QY 268 EGF-ESPOGKITSCTEK-----GTWRESLTCTTILKINDVSL 306  
Db 968 EAFCEVGGSLCVACADENQESPMTCQCRVTESDGVDRQPREKCECYNLNDASL 1026

RESULT 14  
LTBL MOUSE  
ID LTBL MOUSE STANDARD; PRT: 1713 AA.  
AC Q8CG19; Q88349; Q8ENW7; Q8C7F5; Q8C1R0;  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Latent transforming growth factor beta binding protein, isoform 1L  
DE precursor (LTBP-1) (Transforming growth factor beta-1 binding protein  
GN LTBP1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=12711388;  
RA Noguera I., Obata H., Gualandris A., Covin P., Rifkin D.B.;  
RT "Molecular cloning of the mouse ltbp-1 gene reveals tissue specific  
RT expression of alternatively spliced forms.";  
RL Gene 308:31-41(2003).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=129/SvJ;  
RA Weiskirchen R., Moser M., Guenther K., Weiskirchen S., Gressner A.M.;  
RT "The murine latent transforming growth factor-beta binding protein  
RT (ltbp-1) is alternatively spliced, and maps to a region syntenic to  
RT human chromosome 2p21-22.";  
RL Gene 308:43-52(2003).  
RN [3]  
RP SEQUENCE OF 1112-1713 FROM N.A.  
RC STRAIN=C57BL/6J;  
RC TISSUE=Aorta, Liver, and Vein;  
RX PubMed=12466851;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikaide I., Osato N., Saito P., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojodori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Butalov S., Heisel K.W.,  
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Guatinich S., Hirokawa N., Jackson J.J., Jarvis E.D.,  
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Sample C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wallestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
CC -1- SUBUNIT: The large latent complex of TGF-beta1 from platelets is  
CC composed of the TGF-beta1 molecule noncovalently associated with a  
CC disulfide-bonded complex of a dimer of the N-terminal propeptide  
CC of the TGF-beta1 precursor and a third component denoted TGF-  
CC beta1-BP. TGF-beta1-BP does not bind directly to active TGF-BETA1.  
CC Binds to fibrillin (By similarity).  
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1L;  
CC IsoId=Q8CG19-1; Sequence=Displayed;  
CC Name=1S;  
CC IsoId=Q8CG18-1; Sequence=External;  
CC -1- PTM: Contains hydroxylated asparagine residues (By similarity).  
CC -1- SIMILARITY: Contains 16 EGF-like domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF346465; AAN77250.1; JOINED.  
CC EMBL; AF346434; AAN77250.1; JOINED.  
CC EMBL; AF346435; AAN77250.1; JOINED.  
CC EMBL; AF346436; AAN77250.1; JOINED.  
CC EMBL; AF346437; AAN77250.1; JOINED.  
CC EMBL; AF346438; AAN77250.1; JOINED.  
CC EMBL; AF346439; AAN77250.1; JOINED.  
CC EMBL; AF346440; AAN77250.1; JOINED.  
CC EMBL; AF346441; AAN77250.1; JOINED.  
CC EMBL; AF346442; AAN77250.1; JOINED.  
CC EMBL; AF346443; AAN77250.1; JOINED.  
CC EMBL; AF346444; AAN77250.1; JOINED.  
CC EMBL; AF346445; AAN77250.1; JOINED.  
CC EMBL; AF346446; AAN77250.1; JOINED.  
CC EMBL; AF346447; AAN77250.1; JOINED.  
CC EMBL; AF346448; AAN77250.1; JOINED.  
CC EMBL; AF346449; AAN77250.1; JOINED.  
CC EMBL; AF346450; AAN77250.1; JOINED.  
CC EMBL; AF346451; AAN77250.1; JOINED.  
CC EMBL; AF346452; AAN77250.1; JOINED.  
CC EMBL; AF346453; AAN77250.1; JOINED.  
CC EMBL; AF346454; AAN77250.1; JOINED.  
CC EMBL; AF346455; AAN77250.1; JOINED.  
CC EMBL; AF346456; AAN77250.1; JOINED.  
CC EMBL; AF346457; AAN77250.1; JOINED.  
CC EMBL; AF346458; AAN77250.1; JOINED.  
CC EMBL; AF346459; AAN77250.1; JOINED.  
CC EMBL; AF346460; AAN77250.1; JOINED.  
CC EMBL; AF346461; AAN77250.1; JOINED.  
CC EMBL; AF346462; AAN77250.1; JOINED.  
CC EMBL; AF346463; AAN77250.1; JOINED.  
CC EMBL; AF346464; AAN77250.1; JOINED.  
CC EMBL; AF022889; AAC33307.1; --  
CC EMBL; AF143161; AAN38831.1; ALT\_SEQ.  
CC EMBL; AK050380; BAC34222.1; --  
CC EMBL; AK080024; BAC37808.1; --  
CC HSPD; P08709; 1BF9.  
CC MGD; MG1:109151; Ltbp1.  
CC InterPro; IPR000152; Asx hydroxyl.  
CC InterPro; IPR001881; EGF\_Ca.  
CC InterPro; IPR006209; EGF\_Like.  
CC InterPro; IPR002212; Fibrin-assoc.  
CC InterPro; IPR006210; IEGF.  
CC Pfam; PF00608; EGF; 16.  
CC Pfam; PF00683; TB; 4.

DR SMART; SM00181; EGF; 18.  
 DR SMART; SM00179; EGF CA; 16.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 13.  
 DR PROSITE; PS00022; EGF 1; 2.  
 DR PROSITE; PS01186; EGF 2; 11.  
 DR PROSITE; PS01187; EGF CA; 15.  
 DR Growth factor binding; Repeat; EGF-like domain;  
 DR Hydroxylation; Signal; Glycoprotein;  
 DR Alternative splicing.  
 FT SIGNAL 1 23  
 FT CHAIN 24 1713  
 FT  
 FT DOMAIN 619 659  
 FT REPEAT 667 732  
 FT DOMAIN 866 907  
 FT DOMAIN 908 949  
 FT DOMAIN 950 990  
 FT DOMAIN 991 1030  
 FT DOMAIN 1031 1071  
 FT DOMAIN 1072 1112  
 FT DOMAIN 1113 1153  
 FT DOMAIN 1154 1194  
 FT DOMAIN 1195 1236  
 FT DOMAIN 1237 1278  
 FT DOMAIN 1279 1321  
 FT REPEAT 1336 1403  
 FT DOMAIN 1416 1458  
 FT DOMAIN 1459 1499  
 FT REPEAT 1509 1581  
 FT DOMAIN 1613 1653  
 FT DOMAIN 1654 1698  
 FT DISULFID 623 634  
 FT DISULFID 629 643  
 FT DISULFID 645 658  
 FT DISULFID 870 882  
 FT DISULFID 877 891  
 FT DISULFID 893 906  
 FT DISULFID 912 924  
 FT DISULFID 919 933  
 FT DISULFID 935 948  
 FT DISULFID 954 965  
 FT DISULFID 960 974  
 FT DISULFID 977 989  
 FT DISULFID 995 1006  
 FT DISULFID 1001 1015  
 FT DISULFID 1018 1029  
 FT DISULFID 1035 1046  
 FT DISULFID 1041 1055  
 FT DISULFID 1057 1070  
 FT DISULFID 1076 1087  
 FT DISULFID 1082 1096  
 FT DISULFID 1098 1111  
 FT DISULFID 1117 1128  
 FT DISULFID 1133 1137  
 FT DISULFID 1139 1152  
 FT DISULFID 1158 1170  
 FT DISULFID 1165 1179  
 FT DISULFID 1181 1193  
 FT DISULFID 1199 1211  
 FT DISULFID 1205 1220  
 FT DISULFID 1222 1235  
 FT DISULFID 1241 1253  
 FT DISULFID 1247 1262  
 FT DISULFID 1264 1277  
 FT DISULFID 1283 1295  
 FT DISULFID 1290 1304  
 FT DISULFID 1306 1320  
 FT DISULFID 1420 1433  
 Query Match 6.4%; Score 256; DB 1; Length 1713;  
 Best Local Similarity 26.4%; Pred. No. 2.2e-09;  
 Matches 79; Conservative 31; Mismatches 103; Indels 86; Gaps 14;

34 DGLDVCATCH--BHATCOOREBKIKICINYP-----VNGRTQCVDKNECQFATLVGN 87  
 1112 EIDDECEHHHLCSHGQCRNTEGSCFQVCNQGRASVLGD---HCEIDINEEDSSVCQGG 1168  
 88 HTSCHNTPGFYCICLEGYRATNNKTFIPNDGTCTDIDECEVSLCRHGRGVNTHGS 147  
 1169 --DCINTAGSYDCTCPDGFQL-----NDNKGCDLINECAQPLCGSHCECLNTQGS 1217  
 148 FFCYCWGKYLPRNGPEPFHTTDAISCTEIDCGTTPPEVDGVIIGNYTSLSGSQVRYACR 207  
 1218 FHCVCQSG-----FSSADGRTCEDIDCVNNTVCDSH---GFCNDTAGSFRCILCY 1265  
 208 EGFSPFVPEDTVSCTGLGTWESPKLHCQBCINCGNPPMEHAILVGNHSSRLGGVARYVCO 267  
 1266 QG-FQAPQD-----GGQ-----CVDV-----NECELLSG-----VCG 1291  
 268 EGF-ESFGKINSVCTEK-----GTWRESTUTCTEILTINKINDVSL 306  
 1292 EAFCEVVEGSLFCVCADENGEYSMTQCRSRVTDGSDVQRPREEKCEKYNLNDASL 1350

RESULT 15  
 LTBL RAT  
 ID LTBL RAT STANDARD; PRT; 1712 AA.  
 AC Q00918;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Latent transforming growth factor beta binding protein 1 precursor  
 DE (LTBP-1) (Transforming growth factor beta-1 binding protein 1) (TGF-  
 DE beta1-BP-1) (Transforming growth factor beta-1 masking protein, large  
 DE subunit).  
 DE LTBP1.  
 GN Rattus norvegicus (Rat).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=91062373; PubMed=2247454;  
 RA Tsuji T., Okada F., Yamaguchi K., Nakamura T.;  
 RT "Molecular cloning of the large subunit of transforming growth factor  
 RT type beta masking protein and expression of the mRNA in various rat  
 RT tissues.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8835-8839(1990).  
 CC -1- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETAL FROM PLATELETS IS  
 CC COMPOSED OF THE TGF-BETAL MOLECULE NONCOVALENTLY ASSOCIATED WITH  
 CC A MASKING PROTEIN CONSISTING OF A DISULFIDE-BONDED COMPLEX OF A  
 CC DIMER OF THE N-TERMINAL PROPEPTIDE OF THE TGF-BETAL PRECURSOR AND  
 CC A THIRD COMPONENT DENOTED TGF-BETAL-BP (OR MP LARGE SUBUNIT).  
 CC TGF-BETAL-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETAL.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Contains 18 EGF-like domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M55431; AAA42235.1; -  
 CC PIR; A38261; A38261.  
 CC HSSP; P16109; 1FSS.  
 CC InterPro; IPR000152; Asx\_hydroxyl.  
 CC InterPro; IPR001881; EGF CA.  
 CC InterPro; IPR006209; EGF like.  
 CC InterPro; IPR002212; Fibril-assoc.  
 CC Pfam; PF00008; EGF; 16.  
 CC Pfam; PF00683; TB; 4.  
 CC SMART; SM00179; EGF CA; 13.  
 CC PROSITE; PS00010; ASX\_HYDROXYL; 13.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2004, 17:35:09 ; Search time 21 Seconds

(without alignments)  
3420.857 Million cell updates/sec

Title: US-10-066-269-58

Perfect score: 4029

Sequence: 1 MGRGPWDAGPSRRLLPLLLL.....VGLSLAWIILFLSFAV 747

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3002	74.5	570	T46261	hypothetical prote
2	294.5	7.3	2871	A55624	fibrillin-1 precu
3	290.5	7.2	3002	A47221	fibrillin 1 - bovi
4	283.5	7.0	2871	A55567	hypothetical prote
5	283	7.0	3507	T34513	fibrillin-2 precu
6	281	7.0	2907	A57278	fibrillin-2 precu
7	278.5	6.9	2918	A54105	uromodulin precurs
8	277.5	6.9	640	A30452	probable hormone r
9	274.5	6.8	886	A57172	hypothetical prote
10	274	6.8	1106	T18739	latent transformin
11	272.5	6.8	1820	A55494	uromodulin precurs
12	267	6.6	642	S52111	transforming growt
13	260.5	6.5	1394	A35626	masking protein pr
14	251	6.2	1712	A38261	nidogen precursor
15	246.5	6.1	1247	MMHUND	fibulin, splice fo
16	244.5	6.1	705	S34968	fibulin, splice fo
17	243.5	6.0	685	S78040	notch 3 protein -
18	241.5	6.0	2318	S45306	fibulin-2 precurs
19	240	6.0	1221	A49457	tenascin precursor
20	239.5	5.9	2019	J01322	extracellular prot
21	236	5.9	387	T38449	EGF repeat transme
22	236	5.9	1687	T30176	uromodulin precurs
23	235	5.8	644	A40212	Tamm-Horsfall prot
24	235	5.8	644	T184634	tenascin-C - human
25	234	5.8	2201	A32160	fibropellin C prec
26	233	5.8	570	A48836	hypothetical prote
27	232	5.8	1620	T27283	notch3 protein - h
28	231.5	5.7	2321	S78549	nidogen precursor
29	231	5.7	1245	MMMSND	

ALIGNMENTS

RESULT 1

T46261

hypothetical protein DKFp761E1824.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000

R;Bloeker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23032

A;Accession: T46261

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-570 <AAA>

A;Cross-references: EMBL:AL137432

A;Experimental source: adult amygdala; clone DKFp761E1824

C;Genetics:

A;Note: DKFp761E1824.1

Query Match 74.5%; Score 3002; DB 2; Length 570;  
Best Local Similarity 100.0%; Pred. No. 1.4e-181; Indels 0; Gaps 0;  
Matches 570; Conservative 0; Mismatches 0

Qy	178	DCGTPPEVDPGVIIGNYTSLSGQVRYACREGFFSVPEPTVSSCTGLGTWESPKLHCQEI	237
Db	1	DCGTPPEVDPGVIIGNYTSLSGQVRYACREGFFSVPEPTVSSCTGLGTWESPKLHCQEI	60
Qy	238	NCGNPPEMHAILVGNHSSRLGVARYVCOEGFSPGGKITSVCTEKGTWRESTLTCTEI	297
Db	61	NCGNPPEMHAILVGNHSSRLGVARYVCOEGFSPGGKITSVCTEKGTWRESTLTCTEI	120
Qy	298	LTKINDVSLFNDTCVRWQINRRINPKISYVISIKQRLDPMESVREETVNLTTDSRTPE	357
Db	121	LTKINDVSLFNDTCVRWQINRRINPKISYVISIKQRLDPMESVREETVNLTTDSRTPE	180
Qy	358	VCALYPGNGNYTNISTAPRSMFAVIGQFAEVDLLEDGGSFNISINETCLKNRNS	417
Db	181	VCALYPGNGNYTNISTAPRSMFAVIGQFAEVDLLEDGGSFNISINETCLKNRNS	240
Qy	418	RKVGSHMYQFTVLGQWYLANFSHATSNFTTREQVPVVCCLDLYPTTDTYVNVTLRSP	477
Db	241	RKVGSHMYQFTVLGQWYLANFSHATSNFTTREQVPVVCCLDLYPTTDTYVNVTLRSP	300
Qy	478	KRHSVQITTIATPPAVKQTIISNIGSFNETCLWRNSIKTADMEWYLPHIIGQWYQKEPAQ	537
Db	301	KRHSVQITTIATPPAVKQTIISNIGSFNETCLWRNSIKTADMEWYLPHIIGQWYQKEPAQ	360
Qy	538	EMTFNIISSSRDPEVCLDRPGTNVNSLRALSSSELPPVISLTQTTEPPLPEVEFTVH	597
Db	361	EMTFNIISSSRDPEVCLDRPGTNVNSLRALSSSELPPVISLTQTTEPPLPEVEFTVH	420
Qy	598	RGPPLRLRKAKENKNGPISSVQLVPLALOSTFSCDSEGAFFSNASDADYVAEL	657

Db 421 RGPLRLRLKAKKNGKPISSQVQLVLPPLA:QSTFSCDSGASFFSNASDADGYVAAEL 480  
QY 658 LAKVDPDDAMEIPGDRLYGCEYNADPKRSDYCIILRTITSEWKKVRRHSCAVNAQVXD 717  
Db 481 LAKVDPDDAMEIPGDRLYGCEYNADPKRSDYCIILRTITSEWKKVRRHSCAVNAQVXD 540  
QY 718 SSLMLQWAGVGLGSLAVVILTLFSAV 747  
Db 541 SSLMLQWAGVGLGSLAVVILTLFSAV 570

RESULT 2  
A5624  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 02-Aug-2002  
C:Accession: A5624  
R.Yin, W.J. Smiley, E.; Germiller, J.; Sanguinetti, C.; Lawton, T.; Pereira, L.; Ramirez, J. Biol. Chem. 270, 1798-1806, 1995  
A:Title: Primary structure and developmental expression of Fbn-1, the mouse fibrillin gene  
A:Reference number: A5624; MUID:95130561; PMID:7829516  
A:Accession: A5624  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2871 <YIN>  
A:Cross-references: GB:L29454; NID:G575509; PIDN:AAA56840.1; PID:G575510  
C:Genetics:  
A:Gene: Fbn-1  
C:Superfamily: fibrillin 1; EGF homology  
F:1201-1236/Domain: EGF homology <EGF>

Query Match 7.3%; Score 294.4; DB 2; Length 2871;  
Best Local Similarity 26.4%; Pred. No. 3.4e-10;  
Matches 83; Conservative 33; Mismatches 94; Indels 101; Gaps 15;

QY 37 DVACATCHEATCOOREGKKICINYGFGV-NGRTQCVDNKCEQFGATVCGNHTSCHNTP 95  
Db 1289 NICLS---GFCENTKSGFICHDMGYSKKGKTGCTDINECEIGAH-NCGRHAVCTNTA 1343  
QY 96 GGFYICILEGYRATNNKTFIPNDGTFCTDIDEC-EVSGLCRHGGRCVNTGHSFECYCMD 154  
Db 1344 GSFKCSGPGW-----IGDGKICTDIDECNTHGTHMCSHADCKNTGMSYRCLCD 1393  
QY 155 GYLPRNGPEPHPTTATSCTEID-----CGTPP--EVPDGYIIGNYTSLSGSQVRY 204  
Db 1394 GY-----FGDGFTCTDIDECSENILNCGNQCCLNAPGY-----RC 1429  
QY 205 ACREGFTSVPTDVTSSCTGLGTWESPKLHCQKINCGNPPPMRHAILVGNHSSRLGQVARY 264  
Db 1430 ECDMGF--VPSADGKACEDIDCSLPNI-CVFGTCHNLP-----GLFRC 1470  
QY 265 VCQSGFE--SPGKITSVCTEKGTWRESTVCTEILTKINDVSLFNDTCVR-----313  
Db 1471 ECEIGYELDRSGG-----NCTDVNECLDPTTCTSGNCVNTFGSYTDC 1513  
QY 314 ---WQINSRI 321  
Db 1514 SPDFELNPTV 1524

RESULT 3  
A4721  
fibrillin 1 precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 02-Jun-1995 #sequence\_revision 25-Apr-1997 #text\_change 02-Aug-2002  
C:Accession: A4721; I54355; S17064; I59574; S6211; A34198  
R:Corson, G.M.; Chabberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.  
Genomics 17, 476-484, 1993  
A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure  
A:Reference number: A4721; MUID:94010947; PMID:7691719  
A:Accession: A4721  
A:Molecule type: mRNA  
A:Residues: 1-337, 'T', 339-1029 <COR>

A:Cross-references: GB:X63556  
R:Pereira, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Pangillan, T.; Bona  
Hum. Mol. Genet. 2, 961-968, 1993  
A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene  
A:Reference number: I54355; MUID:93372860; PMID:8364578  
A:Accession: I54355  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 132-3002 <PER>  
A:Cross-references: GB:LI3923; NID:G306745; PIDN:AA02036.1; PID:G306746  
R:Masien, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.  
Nature 352, 334-337, 1991  
A:Title: Partial sequence of a candidate gene for the Marfan syndrome.  
A:Reference number: S17064; MUID:91304568; PMID:1852207  
A:Accession: S17064  
A:Molecule type: mRNA  
A:Residues: 1030-3002 <MAS>  
A:Cross-references: EMBL:X63556  
R:Dietz, H.C.; Valle, D.; Francomano, C.A.; Kendzior, R.J.  
Science 259, 680-683, 1993  
A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.  
A:Reference number: I59574; MUID:93157831; PMID:8430317  
A:Accession: I59574  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 2217-2289, 'T', 2290-2325 <RES>  
A:Cross-references: GB:S54426; NID:G264860; PIDN:AA025244.1; PID:G264861  
R:Ries, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, P.  
Nature 352, 330-334, 1991  
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differ  
A:Reference number: S17062; MUID:91304567; PMID:1852206  
A:Accession: S17062  
A:Molecule type: mRNA  
A:Residues: 'VLVTVVRFISYNKML', 944-1444 <LEE1>  
A:Cross-references: EMBL:X62008; NID:G31398; PIDN:CAB56534.1; PID:G5924015  
A:Accession: S62111  
A:Molecule type: protein  
A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEE2>  
R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.  
J. Biol. Chem. 264, 21381-21385, 1989  
A:Title: Connective tissue microfibrils. Isolation and characterization of three large  
A:Reference number: A34198; MUID:90078246; PMID:2512293  
A:Accession: A34198  
A:Molecule type: protein  
A:Residues: 585-575, 1890-1892, 'I', 1894-1900 <VAD>  
C:Comment: Fibrillin is a major component of elastin-associated microfibrils.  
C:Genetics:  
A:Gene: GDB:FBN1  
A:Cross-references: GDB:127115; OMIM:134797; OMIM:154700  
A:Map position: 15q21.1-15q21.1  
A:Introns: 2236/1; 2258/1; 2297/1  
C:Superfamily: fibrillin 1; EGF homology  
C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein;  
F:1-3002/Product: fibrillin (5'-region exon A splice form) #status predicted  
F:132-3002/Product: fibrillin (5'-region exon C splice form) #status predicted <MARC>  
F:1457-1492/Domain: EGF homology <EGF>  
F:1457-1492/Domain: EGF homology <EGF2>  
F:2262-2295/Domain: EGF homology <EGF1>

Query Match 7.2%; Score 290.5; DB 2; Length 3002;  
Best Local Similarity 26.4%; Pred. No. 6.5e-10;  
Matches 82; Conservative 34; Mismatches 94; Indels 101; Gaps 15;

QY 37 DVACATCHEATCOOREGKKICINYGFGV-NGRTQCVDNKCEQFGATVCGNHTSCHNTP 95  
Db 1420 NICLS---GTCENTKSGFICHDMGYSKKGKTGCTDINECEIGAH-NCGRHAVCTNTA 1474  
QY 96 GGFYICILEGYRATNNKTFIPNDGTFCTDIDEC-EVSGLCRHGGRCVNTGHSFECYCMD 154  
Db 1475 GSFKCSGPGW-----IGDGKICTDIDECNTHGTHMCSHADCKNTGMSYRCLCKE 1524  
QY 155 GYLPRNGPEPHPTTATSCTEID-----CGTPP--EVPDGYIIGNYTSLSGSQVRY 204



Db 1525 GY-----TGDTCTDLDECSNLCNGQCLNAPGGY-----RC 1560  
Qy 205 ACRGFFSVPTVSSCTGLTWSPKHLHCGEINCGNPPPMRHALLVGNHSSRLGGVARY 264  
Db 1561 ECDMGF--VPSADGKACEDIDECSLPNI-CVFGTCHNLP-----GLFRC 1601  
Qy 265 VCQEGFE--SPGGKITSVCTEKGTRWRETLTATLTKINDVSLFNDTCVR-----313  
Db 1602 ECEIGYELDRSG-----NCTDVNECLDPTTCISGNCVNTPGSYICDC 1644  
Qy 314 ---WQINSRRI 321  
Db 1645 PPDFELNPTRV 1655

RESULT 4  
A:Species: Bos primigenius taurus (cattle)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Aug-2002  
C:Accession: A55567  
R:Tilstra, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byers, P.H.  
Genomics 23, 480-485, 1994  
A:Title: Sequence of the coding region of the bovine fibrillin cDNA and localization to  
A:Reference number: A55567; MUID:95137597; PMID:7835900  
A:Accession: A55567  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2871 <full>  
A:Cross-references: GB:128748; NID:G508427; PID:AAA74122.1; PID:G508428  
C:Superfamily: fibrillin 1; EGF homology  
F:1201-1236/Domain: EGF homology <EGF>

Query Match 7.0%; Score 283.5; DB 2; Length 2871;  
Best Local Similarity 26.0%; Pred. No. 1.7e-09;  
Matches 81; Conservative 34; Mismatches 95; Indels 101; Gaps 15;

Qy 37 DVATCHEHATCOOREKKKICINYGFGV-NRTQCVDRKCEQFGATLVGNHTSCHNTP 95  
Db 1289 NICLS---GTCENTKGSFICHCDMGYSKKTKGTCTDINECEIGAH-NCDRHACVNTA 1343  
Qy 96 GGFYICILEGYRATNNKTFIPNDGTFTCTDIDEC-EVSGLCRHGGRCVNTHSGFEYCMD 154  
Db 1344 GSFKCSGPGW-----IGDGIKCTDLDECSNTHMCSQHADCNTWGSYRCLCKE 1393  
Qy 155 GYLPRNGPEPPHPTTDTATCTEID-----CGTTP--EVPDGYIIGNYTSLSGVARY 204  
Db 1394 GY-----TGDTCTDLDECSNLCNGQCLNAPGGY-----RC 1429  
Qy 205 ACRGFFSVPTVSSCTGLTWSPKHLHCGEINCGNPPPMRHALLVGNHSSRLGGVARY 264  
Db 1430 ECDMGF--VPSADGKACEDIDECSLPNI-CVFGTCHNLP-----GLFRC 1470  
Qy 265 VCQEGFE--SPGGKITSVCTEKGTRWRETLTATLTKINDVSLFNDTCVR-----313  
Db 1471 ECEIGYELDRSG-----NCTDVNECLDPTTCISGNCVNTPGSYICDC 1513  
Qy 314 ---WQINSRRI 321  
Db 1514 PPDFELNPTRV 1524

RESULT 5  
T34513  
Hypothetical protein ZK783.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34513  
R:Favell, A.; Vaudin, M.  
submitted to the EMBL Data Library, August 1994  
A:Description: The sequence of C. elegans cosmid ZK783.  
A:Reference number: Z21536  
A:Accession: T34513

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-3507 <FAV>  
A:Cross-references: EMBL:U13646; PID:AA24418.1; GSPDB:GNG0021; CESP:ZK783.1  
A:Experimental source: strain Bristol N2; clone ZK783  
C:Genetics:  
A:Gene: CESP:ZK783.1  
A:Map position: 3  
A:Introns: 14/1, 48/2, 84/1, 196/3, 303/1, 381/1, 586/1, 605/1, 1175/3, 1207/1, 1409/2, 3504/1

Query Match 7.0%; Score 283; DB 2; Length 3507;

Best Local Similarity 21.8%; Pred. No. 2.3e-09;  
Matches 185; Conservative 81; Mismatches 303; Indels 280; Gaps 40;

Qy 22 GLAR---GAAGAPGPDG---LDVCAT---CHEHATCOOREKKKICINYGFGVNGRTQC 71  
Db 1431 GVCRCPEGFEAPPKSCVDVDECATGDNCHESARCONVGVYACFCPTGFRKADGSC 1490  
Qy 72 VDNKFCQFGATLVGNHTSCHNTPG-----GFYCI-----101  
Db 1491 QDIDECHEHATCOOREKKKICINYGFGVNGRTQC-----1550  
Qy 102 -----CLEGYR-----ATNNKTF--115  
Db 1551 SESNMSCVDVTVDGVECKECCMGYKSGKVCEDINECAKAPCSLNANCVNMGTFSC 1610  
Qy 116 -----IPNDGTFTCTDIDECVSGLCRHGGRCVNTHSGFEYCMDGYLPRNGPEPPHPTT 170  
Db 1611 SCKQGVGGDFMCTDINECDERHPCHPAECTNLGSGFKCHSGFEGDGIGKCTNPL--1668  
Qy 171 ATSCHEID--CGTPEVDPGIIIGNYTSLSGVARYACREGFSPEDTVSSCTGLTWE 228  
Db 1669 ERSCEDEKFCGRVDHV-SCLSVRIYNGSLSVCE--CEGFRPEKES--NSCVDIDEC 1723  
Qy 229 SPKLHCQECINCGNPPMRHALLVGNHSSRLGGVARYCQEGFSPGKITSVCTSKGTWR 288  
Db 1724 ESRNNDPAS-----AVCVNTE---GSYRCECAEGYEGGG---VCTDIDEC 1765  
Qy 289 ESTLCTTELTAKINDVSLF-----NDTCVRQWINSRINPKISYVLSIKGRLD 337  
Db 1766 RGMAGCDMSAMCINRMGSGCKMAGYTGDACTKIE-----EBPK-----SDKTACTD 1815  
Qy 338 PMESVRE-ETVNLITDSTRPEVCLALYPGNTYV-----I 372  
Db 1816 EWSRICELEKXCTVDEEVPCGACLPG-HHPINGTCOSLOISGLCAKNDCKNHAECI 1874  
Qy 373 STAPPRSRMPAVIGF-----QTAEDVLLDD-----GSFN-----ISIFN 407  
Db 1875 DIHPDSHFCSGPDGFIGDMICDDVDENACMCDDEKTCENTIGSFNCVCLGFKKVD 1934  
Qy 408 ETCL-----KLNRRSRKVGSEH-----MYQFTVLGQWYLANFSAHSFNFTREQVP 455  
Db 1935 EKVDEKQPKREKIEIDENSSSSNSGQKPTTKG---IVSSTSATSESTAE--P 1988  
Qy 456 VVCLDLYPTDVTYVNTLLRSRKR-----HSVQITATPPAVKQIISNIS---500  
Db 1989 HVTTISSTTS-TKDMTSSKSPENVMTSSESPVSTSSKSTTASSETTVSTSPSSSE 2047  
Qy 501 -----GFNEICLRWSIKTADMBEMYLFIHWGQWYQKEPAQWTFNISSSRDPEV-552  
Db 2048 APLTSSPATTEVTESVKTTPKE-----ESSSEITVLSKSKS--PEVT 2091  
Qy 553 --CLDLRPGTNYNS-----LRALSSSELPWISLTQI-----TPPLP--589  
Db 2092 ESSVSKSPSTPSTTSQSVSTVTPETSKSTVLSSEAPVTSTSTVHTSTSTPSTSLASST 2151  
Qy 590 --EVEFTVHRGRLPLRLRKAKENGPISVQVVLPLALOSTFSCDEGASSPFSNAS 647  
Db 2152 TGDNTSTTSTSLASVKSTSAPE--GTSAS-----VAPVKL--SSLSFDVDSQSTKTFTAT 2204  
Qy 648 DADGYVAEE 656  
:: ::

Db 2205 ESSTVQASE 2213

## RESULT 6

A57278  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Feb-1996 #sequence\_revision 23-Feb-1996 #text\_change 02-Aug-2002  
C:Accession: A57278  
R:Zhang, H.; Hu, W.; Ramirez, F.  
J. Cell Biol. 129, 1165-1176, 1995  
A:Title: Developmental expression of fibrillin genes suggests heterogeneity of extracellular matrix components  
A:Reference number: A57278; MUID:95263670; PMID:7744963  
A:Accession: A57278  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2907 <ZHA>  
A:Cross-references: GB:139790; NID:9762830; PID:AAA74908.1; PID:g762831  
C:Superfamily: fibrillin 1; EGF homology  
F:1259-1274/Domain: EGF homology <EGF>  
F:2498-2523/Domain: EGF homology <EGF>

Query Match 7.0%; Score 281; DB 2; Length 2907;  
Best Local Similarity 28.8%; Pred. No. 3.5e-09;  
Matches 83; Conservative 39; Mismatches 114; Indels 52; Gaps 15;  
QY 25 RGAAGAPDGLDVA-TCHHATCOOREGKIKICNYGVNGRTQCVKNECPGATL 83  
DB 1353 KGTGCTDVTDECEIAGNCDHMSCLNVPGRFKSCREGVWNG-IKCIDLDECANG-TH 1410  
QY 84 VCGNHTSCHNTPGGFYCICLEGVATNNKTFIPNDGTFTDIDEC-EVSGLCRGGRCV 142  
DB 1411 QCSINACQVNTPGSYRCACSEGF-----TGDFTCSDVDECAENTNLCN-GQL 1459  
QY 143 NTHSGFPCYCMGVLPRNGPEPHPTDATTCTEID-CGTPPVPGYIIGNYTSLSGQ 201  
DB 1460 NVPGAYCECEMG-----FTPASDSRQCQDIDECSP-----QNICVFQTCNNLP 1505  
QY 202 VRYACREGFFVPEDTVSSCTGLTWESPKLHC-----QEINGCNPPMR-HAI 249  
DB 1506 FHCICDDGY--ELDRTGNCCTDIDECADP-INCWGLCVNTPGRYECNC--PPDFQLNAT 1560  
QY 250 LVGNHSSRLGGVARYVQEGFSPGKITSCTEKGRTWRESTUTCTEI 297  
DB 1561 GVGCVDRNVRGN-----CYLFGPRGDSGLS-NTAEAGVGSRSRSCCCSL 1603

## RESULT 7

A54105  
fibrillin-2 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 02-Aug-2002  
C:Accession: A54105; S17063; S31101  
R:Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sarguineti, C.; Bonadio, J.; Mecham, J. Cell Biol. 124, 885-863, 1994  
A:Title: Structure and expression of fibrillin-2, a novel microfibrillar component precursor  
A:Reference number: A54105; MUID:94165150; PMID:8120105  
A:Accession: A54105  
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-2918 <ZHA>  
A:Cross-references: GB:U03272  
R:Lee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, P. Nature 352, 330-334, 1991  
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two different genes  
A:Reference number: S17062; MUID:91304567; PMID:1852206  
A:Accession: S17063  
A:Molecule type: mRNA  
A:Residues: 752-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>  
A:Cross-references: EMBL:X62009  
R:Milwicz, D.M.  
submitted to the EMBL Data Library, December 1992  
A:Reference number: S31101

A:Accession: S31101  
A:Molecule type: mRNA  
A:Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 19;  
A:Cross-references: EMBL:X62009  
C:Genetics:  
A:Gene: GDB:FBN2  
A:Cross-references: GDB:128122; OMIM:121050  
A:Map position: 5q23-q31  
C:Superfamily: fibrillin 1; EGF homology  
C:Keywords: extracellular protein  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-2918/Product: fibrillin-2 #status predicted <MAT>  
F:1245-1280/Domain: EGF homology <EGF>  
F:1290-2013/Domain: EGF homology <EGF>

Query Match 6.9%; Score 278.5; DB 2; Length 2918;  
Best Local Similarity 28.8%; Pred. No. 3.6e-09;  
Matches 90; Conservative 30; Mismatches 126; Indels 67; Gaps 17;  
QY 2 GRGFWDAGPSRRLLPLLLGLLARG-----AAGAFGP--DGLDV-----CATCHEHAT-- 47  
DB 449 GYGPGGTG----FIPPGNGPSPGVGGAGGQGPITGLITLNTIDICKHANLC 504  
QY 48 ----COOREGKIKICNYGVNGRTQCVKNECPGATLVCGNHTSCHNTPGGFYCICL 103  
DB 505 LNERCIPVTSSYRCENMGYKQDAGDCIDVDEC---TSNPCTN-GDCVNTFSSYCKCH 560  
QY 104 EGYRATNNKTFIPNDGTFTDIDECVSLCRHGRVCNTHGSPFCYCMGVLPRNGPE 163  
DB 561 AGFPTPTKQA-----CIDIDECIQNGVLCKGRVCNDSGSPQCICNAG----- 604  
QY 164 PEHPTDATTCTEID-CGTPPEVPDGVIIIGNYTSLSGQVRYACREGFFVPEDTVSSCT 222  
DB 605 -FELTDGKNCVDHDCCTTNMCLNGMCINE-----DGSFKICXPGFVLNPGRY--CT 656  
QY 223 GLGTWESPKL----HC-----QEINGCNPP-----EMRHAILVGNH--SSRLGGVARYVQ 268  
DB 657 DVDECOPTPGICMNGHCINSESGSPRCDCPPLAVGMDGRVCVDTHMRSTCYGIGKKGVCVR 716  
QY 269 GFSPGKITSVC 281  
DB 717 PF--PGNVTKSEC 727

## RESULT 8

A30452  
uromodulin precursor - human  
N:Alternate names: Tamm-Horsfall urinary glycoprotein; uromuclid  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A30452; A30453  
R:Pennica, D.; Kohr, W.J.; Kuang, W.J.; Glaister, D.; Aggarwal, B.B.; Chen, E.Y.; Goedd, Science 236, 83-88, 1987  
A:Title: Identification of human uromodulin as the Tamm-Horsfall urinary glycoprotein.  
A:Reference number: A30452; MUID:87177970; PMID:3453112  
A:Accession: A30452  
A:Molecule type: DNA  
A:Residues: 1-640 <PEN>  
A:Cross-references: GB:M15881; NID:9340163; PIDN:AAA36798.1; PID:9340164  
A:Note: parts of this sequence, including the amino end of the mature protein, were determined by R. Hession, C.; Decker, J.M.; Sherblom, A.P.; Kumar, S.; Yue, C.C.; Mattaliano, R.J.; Tilz Science 237, 1479-1484, 1987  
A:Title: Uromodulin (Tamm-Horsfall glycoprotein): A renal ligand for lymphokines.  
A:Reference number: A30453; MUID:87319675; PMID:3498215  
A:Accession: A30453  
A:Molecule type: mRNA  
A:Residues: 1-640 <HES>  
A:Cross-references: GB:M17778  
A:Note: the authors translated the codon GCG for residue 381 as Asp  
A:Note: parts of this sequence, including the amino end of the mature protein, were determined by R. Rindler, M.J.; Naik, S.S.; Li, N.; Hoops, T.C.; Peraldi, M.N. J. Biol. Chem. 265, 20784-20789, 1990  
A:Title: Uromodulin (Tamm-Horsfall Glycoprotein/Uromuclid) is a phosphatidylinositol-li

A;Reference number: A30454; MUID:91065873; PMID:2249987  
A;Contents: annotation; GPI-anchor  
C;Comment: The protein is anchored to the cell membrane by a phosphatidylinositol linkage  
e membrane-bound form.  
C;Genetics: GDB:UMOD

A;Gene: GDB:UMOD  
A;Cross-references: GDB:133729; OMIM:191845  
A;Map position: 16p13.11-16p12.3  
A;Introns: 30/1; 325/1; 394/3; 444/2; 580/3; 608/1; 621/1  
A;Note: the first intron occurs before the initiator codon  
C;Superfamily: uromodulin; EGF homology; membrane glycoprotein 2 homology; ZP domain hom  
C;Keywords: duplication; glycoprotein; lipoprotein; phosphatidylinositol linkage; phosph  
F;1-24/Domain: signal sequence #status predicted <SIG>  
F;25-640/Product: uromodulin #status predicted <MAT>  
F;69-106/Domain: EGF homology <EGF>  
F;170-639/Domain: ZP domain homology <ZPH>  
F;334-583/Domain: membrane glycoprotein 2 homology <MGH>  
F;625-640/Domain: transmembrane #status predicted <TRM>  
F;38, 76, 80, 275, 322, 513/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;232,396/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 6.9%; Score 277.5; DB 1; Length 640;  
Best Local Similarity 23.4%; Pred. No. 6.5e-10;  
Matches 163; Conservative 77; Mismatches 217; Indels 239; Gaps 38;

```
QY 39 CATCHEATCOQREGKICICNYGVNGRTQCVKNECFGATLVCGNHTSCHNTPGCF 98
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 32 CSECHNATCTEAEVTTCTCGEGTGDGLT-CVDLDECAIPGAHNCNSANSCVNTPEGF 90
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 99 YICLCGVRATNNKTFIFNDGTFCTDIDCEVSGL--CRHGRCVNTGSPFECVMDGY 156
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 91 SCVCPGFRLL-----SPGLGCTDVEDCAPGLSHCHALATCVNVVGYLCPAGY 141
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 157 -----LPNN-----GPPFPHTTD----- 170
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 142 RGDGWHCECPGCGPLDCVPGDALVCADPCQAHRTLDVYRSTEGYACDTRDG 201
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 171 -----ATSCETI-DC-----GPPPEVPDGIIGNYTS-----SLG 199
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 202 WYRFVQGGARMAETCVPLVRCNTAPMWLNTHPSDDEGIYSRKACAWSHGCCLDWAS 261
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 200 SQVRYACREGFF-----SVPEDTVSSCTG---LGWTE-----SPKLHC---QE 236
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 262 VQVK-ACAGGYVYVNLTAPECHLAYCTDPSVEGTCECSIDEDCKSNNGRWHQCKQD 320
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 237 INCNPPMEHAILVGNHSLRGVARYVQ---EGFESPGKI-----TSVCT-----E 283
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 321 FNITDLSLLEHLECGANDMK---VSLGKCKLKSGLFD---KVFMVLSDSRCSGENDRD 373
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 284 KGTWRESTLT-----CTEILTAKINDVSLFNQTC-VRWQINSRRINPKISYV----- 328
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 374 NRDK-VSVVTPARDGPGTGLTRNETHATVSNLYLADEILIRDLNLIKINFACSYPLDMK 432
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 329 ISIKGQRLDPNES---VREBTVNLTDSRTPEVCLALYPGTNTWNISTAPRRSPAVI 385
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 433 VSLK-TALQPMVSLAIRVGGTGMT-----VRMALFQTPSYT-----QP----- 471
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 386 GFOTAEVDLDDGSGFNISFNETCLNRRSRKVGSEHYQFTVLGRWYLANFASHATS 445
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 472 -YQSSVTL-----STEAFIVGTMLDGDLR-----FALLMTWCYATPSSNATD 516
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 446 --FNFTTREGQPVCLDLYPTTYVNVTL--LRSPKHSVQITTIATPPAVKQTISNIG 501
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 517 PLKYFIIDRCP-----HTRDSTIQVNGESSQGRFSVQM----- 552
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 502 FNETCLWRISIKTADME---EMVLFHMGQVYQKEFAQEMTFNTSSSRDPEVCLDLR 557
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 553 -----FRFAGNVDLVYLHCEVTLCDTMMKCK-KPTCSGTRFR-SGVIDOSRVNLG 602
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 558 PGTNYN-----SLRALSELPPVIMTQIT 584
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 603 PITRKGVQATVSRAFSSGLLKWLPILLSATLT 638
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

## RESULT 9

A57172  
probable hormone receptor EMR1 precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 11-Jan-2000  
C;Accession: A57172  
A;Title: EMR1, an unusual member in the family of hormone receptors with seven transmemt  
Genomics 26, 334-344, 1995  
A;Reference number: A57172; MUID:95324926; PMID:7601460  
A;Accession: A57172  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-886 <BAU>  
A;Cross-references: GB:X81479; NID:G784993; PIDN:CAA5732.1; PID:G784994  
C;Genetics:  
A;Gene: GDB:EMR1  
A;Cross-references: GDB:378349; OMIM:600493  
A;Map position: 19p13.3-19p13.3  
C;Superfamily: unassigned EGF-related proteins; EGF homology  
C;Keywords: transmembrane protein  
F;136-170/Domain: EGF homology <EGF>

Query Match 6.8%; Score 274.5; DB 2; Length 886;  
Best Local Similarity 22.8%; Pred. No. 1.5e-09;  
Matches 89; Conservative 53; Mismatches 125; Indels 123; Gaps 17;

```
QY 42 CHEHATCOQREGKIKICNYGVFV-GNGR-----TQCVKNECFGATLVCGNHTSCHN 93
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 41 CPAYATCTNTVDSYCTCKQGLFSSNGQNHFKDPGVKCDIDEGS-QSPQPCGPNSSCKN 99
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 94 TPGGFYICILGVRATNNKTFIFNDGTF-CTDIDCEVSGLCRHGRCVNTGHSFECYC 152
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 100 LSGRYKSCLDGFSPTGNDWVPKPGNFSCDINECLTSRVCEHSDCVNMSGYS CSC 159
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 153 MDGLPRNGPFPFHTTDTATCTEI-DCGTPPEVPD---GYIIGNYTSLSLGSQVRYACR 207
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 160 QVGFISRN-----STCEDVNECADPRACPEHATCNTVGVNYSC-----FCN 200
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 208 EGFSPVEDTVSSCTGLTWESPKLHCQEN-----CGNPPMEHAILVGNHSS 256
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 201 PGFES--SSGHLSCQGL-----KASCEDIDECTEMCPINSTCTNTP----- 239
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 257 RLGGVARYVQCEGPESPGKI-----TSVCT----- 283
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 240 ---GSYFCTCHPGAPSSGQLNFTDQGVCECDIDECRQDPSTCGPNSICTNALGSYSCG 296
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 284 -----KGTWRESTLTCTEILTAKINDVSLFNQTCVWRQINSRRINP-----KISYV 328
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 297 IVGFHPNPEGSQKDNFSCQVLFKCKEDVIPDNKQIQCCQEGTAVRPAYVYVFCQINNI 356
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 329 ISIKGQRLDPNESVREBTVNL--TTDSRTP 356
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 357 FSV-----LDKVCENKTVVSLKNTTESFVP 382
      :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

## RESULT 10

T18739  
hypothetical protein B0393.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T18739  
C;Status: J.  
C;Accession: T18739  
submitted to the EMBL Data Library, September 1994  
A;Reference number: Z19013  
A;Accession: T18739  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1106 <WLL>  
A;Cross-references: EMBL:Z37983; NID:e1519039; PIDN:CAA86058.1; GSPDB:GN00021; CBSP:B039  
A;Experimental source: clone B0393  
C;Genetics:

A:Gene: CESP:B0393.5  
A:Map position: 3  
A:Introns: 15/3; 60/1; 100/3; 343/2; 443/1; 607/3; 609/3; 659/1; 796/1; 876/1; 946/1; 99

Query Match 6.8%; Score 274; DB 2; Length 1106;  
Best Local Similarity 25.0%; Pred. No. 2.1e-09;  
Matches 107; Conservative 53; Mismatches 144; Indels 124; Gaps 24;

QY 39 CATCHEATCOOREGKICICNYGVNGRQCVDKNECF-GATLVCGNHTSCHNTPG 97  
DB 373 CARCDONAKC-----SNGVCTCSEGTGEG-PRCYDVDECEIPGA--VCRDHSICSNITGS 425  
QY 98 FYCICLEGYRATNNKTFIPNDCTFCTDIDHC-EVSLGC---RHGRCVNTHGSEFCYCM 153  
DB 426 FECTCHGGYR-----BDGK-CEDVDHCRCLPKICGDPNKGKCNKDGTFECLCK 475  
QY 154 DOYLFRNGPEFPHFTTATSCTEID-CGTPEVEPDGVIIGNYTSLSQSVRYACREGFFS 212  
DB 476 DGY-----EGDPSSECDRWECKNPDACPNQCNTQ--GGYECE-----CLAGFER 521  
QY 213 VPEDTVSSCTGLTWESPKLHCQIN-CGNPPEMRHAILVGNHSSRLGGVARYVCOEGFE 271  
DB 522 IAEGA-----HCTDRDECAVPCHPAAI-----CSNTRSGYKCECRDGFV 561  
QY 272 SPGGRKITSVCTE-----KGTWRSTLTCTEILTCTEILTKINDVSLFNDTCVRQWINS 318  
DB 562 GDG---KTCHEITLYPISNDSTWIPRSWDSST-----AVPILSEITFGKKYKIYLS 612  
QY 319 RRIIPKISYVISIGQRLDPMS-----VRETVNLTDSRPEVCLALY 363  
DB 613 NGI---ISFDGFLQG-LIDHAETLUKPAIFALHAQFYDREGLVAYTFINDTDAVTLPL 668  
QY 364 PGTNYTNISTAPPERSMPAIVGQTAEDVLEDDGSEFNISIFNETCLNRRSRKVGSE 423  
DB 669 ---MRSSIGIQTA-----MGIENQTKELHIFTFD-----RVQSGSE 703  
QY 424 HMYQETVL 431  
DB 704 NPNSFQIV 711

RESULT 11  
A55494  
latent transforming growth factor-beta-binding protein - human  
C:Species: Homo sapiens (man)  
C:Date: 03-Mar-1995 #sequence\_revision 03-Mar-1995 #text\_change 11-Jan-2000  
C:Accession: A55494  
R:Moran, A.; Olofsson, A.; Stenman, G.; Sahlin, P.; Kanazaki, T.; Claesson-Welsh, L.; ten  
J. Biol. Chem. 269, 32469-32478, 1994  
A:Title: Identification and characterization of LTBP-2, a novel latent transforming grow  
A:Reference number: A55494; MUID:95096101; PMID:7798248  
A:Accession: A55494  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1820 <MOR>  
A:Cross-references: GB:237976  
C:Genetics:  
A:Gene: GDB:LTBP2  
A:Cross-references: GDB:568901  
A:Map position: 11pter-11qter  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
F:1222-1257/Domain: EGF homology <EGF>  
F:1525-1565/Domain: EGF homology <EGF1>

Query Match 6.8%; Score 272.5; DB 2; Length 1820;  
Best Local Similarity 28.1%; Pred. No. 4.8e-09;  
Matches 82; Conservative 25; Mismatches 90; Indels 95; Gaps 16;

QY 22 GLARGAGA-----PGPDGLDVCAATCHE-----HATCOOREGKICICNY 61  
DB 1106 GVCNTNTAGSFCKDCDGGYRPSPLG-DSCEVDDECDPOSSCLGGCKNTVGSYQCLCPQ 1164  
QY 62 GF-VNGRGTQCVDKNECFGATLVCGNHTSCHNTPGFGVCI CLEGYRATNNKTFIPNDG 120

DB 1165 GFQLANG-TVCEDVNECM--GESEHCAPHGECLNSHGSFFCLCAPGVFSAG-----G 1213  
QY 121 TFCCTDIDCEVSLGCRHGRCVNTHGSEFCYCMGYLPRNGPEPFPHTTATSCTEID-- 178  
DB 1214 TSCODVDECATTPC-VGHCVNTGSEFNCLCEGT-----FQSPESGECVVIDEC 1263  
QY 179 -----CGT-----PPEVDPGVIIGNYTSLSQSVRYACREGFFSVPEDTVSSCTGLGT 226  
DB 1264 EDYGDPCVGTWKCEPSP-----GSYRCVLG-----CQGFHMAPNG----- 1299  
QY 227 WESPCLHCQINCGNPEMRHAILVGNHS--SRLLGVARYVCOEGFE-SPGG 275  
DB 1300 -----DCIDIDECANDTMCSEHGFCNDTGDSPRCLDQGFSEISPSG 1340

RESULT 12  
S52111  
uromodulin precursor - mouse  
N:Alternate names: Tamm-Horsfall urinary glycoprotein; uromucoid  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S52111  
R:Prasad, K.; Bates, J.; Badgett, A.; Dell, M.; Sukhatme, V.; Yu, H.; Kumar, S.  
Biochim. Biophys. Acta 1260, 328-332, 1995  
A:Title: Nucleotide sequence and peptide motifs of mouse uromodulin (Tamm-Horsfall prot  
A:Reference number: S52111; MUID:95178555; PMID:7873609  
A:Accession: S52111  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-642 <PRA>  
C:Superfamily: uromodulin; EGF homology; membrane glycoprotein 2 homology; Zp domain ho  
C:Keywords: duplication; glycoprotein; lipoprotein; membrane protein; phosphatidylinosi  
F:69-105/Domain: EGF homology <EG1>  
F:111-147/Domain: EGF homology <EG2>  
F:171-641/Domain: membrane glycoprotein 2 homology <MGH>  
F:335-584/Domain: Zp domain homology <ZPH>

Query Match 6.6%; Score 267; DB 1; Length 642;  
Best Local Similarity 24.2%; Pred. No. 3e-09;  
Matches 99; Conservative 38; Mismatches 140; Indels 132; Gaps 16;

QY 28 AGAPGPDGLDVCAATCHEATCOOREGKICICNYGVNGRQCVDKNECFGATLVCGN 87  
DB 21 AGASNSTEARRCSECHNATCTVDGVVTTCSQTGTGDLV-CEDMDECATPWTNCSN 79  
QY 88 HTSCHNTPGFGYICILEGYRATNNKTFIPNDGTFTCTDIDCEVSLG--CRHGRCVNTH 145  
DB 80 -SSCVNTPGSKFCSCQDGFRLT-----PDLGCTDVDECESEQLSNCHALATCVNTE 129  
QY 146 GSEPCYCMGYLPRNGPEPFPHTTATSCTE-IDCGTPPEVDPGYII-----GN 193  
DB 130 GDYLCVCPKGT---GDGWYCECSPSCPEPGLDC--LPQDPDGKLVQDDPCNTVETL 183  
QY 194 YTSLSQSVRYAC---REGFFSVPEDTVSSCTGLTWESPKLHCQINCGNPEMRHAIL 250  
DB 184 YWRSTGYGVGYSADAGHGWYR-----TGQGVMAETCVPVLACNTAEPN---WL 232  
QY 251 VGNH--SRLLGVARYVCOEGFE-SPGGKITSVCTEKGTRWESLTCTEILTINDVSLND 309  
DB 233 NGSHPSSEGIIVSRKTAH-----WSD 254  
QY 310 TCVRWQINRRINPKISYVISIGQRLDPMSVREETVNLTTDSRTPEVCLALYPTNYT 369  
DB 255 HCCRW-----TDPSSVEGTCECRVDEDCISDNGRWR 315  
QY 370 VNTSTAPPRESMPAIVGQTAEDVLEDDGSEFNISIFNETCLNRRSR 418  
DB 275 YNL-TEPPECNLAYC-----TDPSSVEGTCECRVDEDCISDNGRWR 315

RESULT 13  
A35626

transforming growth factor beta-1-binding protein - human  
C:Species: Homo sapiens (man)  
C>Date: 21-Sep-1990 #sequence\_revision 21-Sep-1990 #text\_change 11-Jan-2000  
C:Accession: A35626  
R:Kanzaki, T.; Olofsson, A.; Moren, A.; Wernstedt, C.; Hellman, U.; Miyazono, K.; Claessens, J.; Kjaer, T. 1051-1061, 1990  
A>Title: TGF-beta1 binding protein: a component of the large latent complex of TGF-beta1  
A:Reference number: A35626; MUID:90275601; PMID:2350783  
A:Status: preliminary; not compared with conceptual translation  
A:Accession: A35626  
A:Molecule type: mRNA  
A:Residues: 1-1394 <KAN>  
A:Cross-references: GB:M34057; NID:G339547; PIDN:AAA61160.1; PID:G339548  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
C:Keywords: alternative splicing  
F:750-791/Domain: EGF homology <EGF>

Query Match 6.5%; Score 260.5; DB 2; Length 1394;  
Best Local Similarity 27.6%; Pred. No. 2a-08;  
Matches 82; Conservative 32; Mismatches 102; Indels 81; Gaps 16;

QY 34 DGLDVCATCH--BHATCQREGKKICICNYGFVNG-RYQVDKNECFQATLVCGNHTS 90  
DB 792 EDIDECQHRHLCAGQCRNTEGFSQVCQDQFRASGLGDHCEIDNELEDKS-VC-QRQD 849  
QY 91 CHNTPGGFYICLEGYRATNNKTFIPNDGTCTDIDCEVSGLCRHGRCVNTGSRPC 150  
DB 850 CINTAGSYDCTCPDGFQ-LDNNKT-----CQDNECEHPGLCPQCECLNTEGSRFC 900  
QY 151 YCMGDLPRNGPFPHTTDTATSCETIDCGTPPEVDGVIIGNYNTSSLSQSVRYACREGF 210  
DB 901 VCQCG-----FSISADGRCTEDIDECVNTVCDSH--GFCNDTAGSFRLCLYCG- 947  
QY 211 FSVPEDEVSSCTGLGTWESPKLHCQECINCNCPENMHAILVGNHSSRLGGVARYVCQGF 270  
DB 948 FQAPQD-----CQG-----CVDV-----NECLLSG-----VCGEAF 974  
QY 271 -BSPGKITSVCTEK-----GTWRESLTCTET-----LTKINDVSL 306  
DB 975 CENVEGSLVCADENQESYPTGQCRSRTSIDLVDVDPQKKEKCYNNLDASL 1031

RESULT 14  
A38261  
masking protein precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 14-Jun-1991 #sequence\_revision 14-Jun-1991 #text\_change 21-Jan-2000  
C:Accession: A38261  
R:Teuji, T.; Okada, F.; Yamaguchi, K.; Nakamura, T.  
Proc. Natl. Acad. Sci. U.S.A. 87, 8835-8839, 1990  
A>Title: Molecular cloning of the large subunit of transforming growth factor type beta  
A:Reference number: A38261; MUID:91062373; PMID:2247454  
A:Accession: A38261  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1712 <TSU>  
A:Cross-references: GB:M55431; NID:G207285; PIDN:AAA42235.1; PID:G207286  
C:Superfamily: unassigned EGF-related proteins; EGF homology; LDL receptor ligand-binding  
F:911-947/Domain: EGF homology <EGF>

Query Match 6.2%; Score 251; DB 2; Length 1712;  
Best Local Similarity 26.4%; Pred. No. 1a-07;  
Matches 79; Conservative 30; Mismatches 104; Indels 86; Gaps 14;

QY 34 DGLDVCATCH--BHATCQREGKKICICNYGFVNG-RYQVDKNECFQATLVCGN 87  
DB 1111 EDIDECQHRHLCAGQCRNTEGFSQVCQDQFRASVLGD--HCEIDNELEDSSVCQGG 1167  
QY 88 HTSCNTPGGFYICLEGYRATNNKTFIPNDGTCTDIDCEVSGLCRHGRCVNTGHS 147  
DB 1168 --DCINTAGSYDCTCPDGLQ-----NUNKGCQDINECAQGLCAPHGECCLNTQGS 1216  
QY 148 PECYCMGDLPRNGPFPHTTDTATSCETIDGTPPEVDPGVIIGNYNTSSLSQSVRYACR 207

DB 1217 FHCVCEQG-----FSISADGRCTEDIDECVNTVCDSH--GFCNDTAGSFRLCLY 1264  
QY 208 EGFESVPEDEVSSCTGLGTWESPKLHCQECINCNCPENMHAILVGNHSSRLGGVARYVCQ 267  
DB 1265 QG-FQAPQD-----CQG-----CVDV-----NECLLSG-----VCQ 1290  
QY 268 EGF-ESPGKITSVCTEK-----GTWRESLTCTETILTKINDVSL 306  
DB 1291 EAFCEVSGSLVCADENQESYPTGQCRSRTSIDLVDVDPQKKEKCYNNLDASL 1349

RESULT 15  
MMHUND  
nidogen precursor - human  
N:Alternate names: entactin  
C:Species: Homo sapiens (man)  
C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 20-Oct-2000  
C:Accession: A33322; A32437; A61367  
R:Nagayoshi, T.; Sanborn, D.; Hickok, N.J.; Olsen, D.R.; Fazio, M.J.; Chu, M.L.; Knowlton, D.A. 581-594, 1989  
A>Title: Human nidogen: complete amino acid sequence and structural domains deduced from DNA  
A:Reference number: A33322; MUID:90091745; PMID:2574658  
A:Accession: A33322  
A:Molecule type: mRNA  
A:Residues: 1-1247 <NAG>  
A:Cross-references: EMBL:M30269  
R:Olsen, D.R.; Nagayoshi, T.; Fazio, M.; Mattei, M.G.; Passage, E.; Weil, D.; Timpl, R.; Am. J. Hum. Genet. 44, 876-885, 1989  
A>Title: Human nidogen: cDNA cloning, cellular expression, and mapping of the gene to chromosome 8  
A:Reference number: A32437; MUID:89270475; PMID:2471408  
A:Accession: A32437  
A:Molecule type: mRNA  
A:Residues: 667-1247 <OLS>  
A:Cross-references: EMBL:M27445; NID:G602466; PIDN:AAA57261.1; PID:G602467  
A:Note: the authors translated the codon AAG for residue 966 as Cys  
R:Fazio, M.J.; O'Leary, J.; Kaehaeri, V.M.; Chen, Y.Q.; Saitta, B.; Uitto, J.  
J. Invest. Dermatol. 97, 281-285, 1991  
A>Title: Human nidogen gene: structural and functional characterization of the 5'-flank  
A:Reference number: A61367; MUID:91302882; PMID:1906509  
A:Accession: A61367  
A:Molecule type: DNA  
A:Residues: 1-28 <FAZ>  
A:Comment: This protein is a basement membrane glycoprotein that forms a complex with la  
C:Genetics:  
A:Gene: GDB:NID  
A:Cross-references: GDB:120236; OMIM:131390  
A:Map position: 1q43-1q43  
C:Superfamily: nidogen; EGF homology; LDL receptor WYTD-containing repeat homology; thyr  
C:Keywords: basement membrane; beta-hydroxyasparagine; calcium binding; cell binding; co  
protein  
F:1-28/Domain: signal sequence #status predicted <SIG>  
F:29-1247/Product: nidogen #status predicted <MAT>  
F:390-425/Domain: EGF homology <EG1>  
F:672-708/Domain: EGF homology <EG2>  
F:702-704/Region: cell attachment (R-G-D) motif  
F:714-750/Domain: EGF homology <EG3>  
F:762-800/Domain: EGF homology <EG4>  
F:806-839/Domain: EGF homology <EG5>  
F:849-919/Domain: thyroglobulin type I repeat homology <THY1>  
F:990-1032/Domain: LDL receptor WYTD-containing repeat homology <YWI>  
F:1033-1075/Domain: LDL receptor WYTD-containing repeat homology <YWI2>  
F:1076-1120/Domain: LDL receptor WYTD-containing repeat homology <YWI3>  
F:1121-1160/Domain: LDL receptor WYTD-containing repeat homology <YWI4>  
F:1161-1197/Domain: LDL receptor WYTD-containing repeat homology <YWI5>  
F:1212-1243/Domain: EGF homology <EG6>  
F:289-296/Binding site: sulfate (Tyr) (covalent) #status predicted  
F:729-819/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted  
F:756/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of laminin) #status predi  
F:1137/Binding site: carbonyl (Asn) (covalent) #status predicted

Query Match 6.1%; Score 246.5; DB 1; Length 1247;  
Best Local Similarity 22.1%; Pred. No. 1.3e-07;

Matches 138; Conservative 56; Mismatches 192; Indels 239; Gaps 30;		
QY	32	GPDLGVCATCHEHATCOOREGK---ICINYGFGVNGRTQCYDKNECOFGATLVCGNH 88
Db	675	GTGCDTNAAC-----RGPRTQFTCE--SIGFRGDGRT-CYDIDECSEQPS-VCGSH 724
QY	89	TSCHTPGGFCICLEGY-----RATNNKTFIPN-----118
Db	725	TICNNHPTFCCEVEGQFSDGTCVAVVQRPINVCETGLHNCIDIPQRAQCIYTGSS 784
QY	119	-----DGTFTDIDCEVSGLCRHGRVCNTHGSPCYCMDGY-----156
Db	785	YTCSLPGFSGDQACQVDECPQS-RCHPFAFCYNTPGSFTCQCKPGYQGDGFRCPGE 843
QY	157	-----LPNGPEPHPTTDA-----TSC-----174
Db	844	VEKTRCQHERHILGAAGATDFQRP---GLFVPECDAGHYAPTQCHGTCYCWCVD 900
QY	175	-----TEIDCG-TTP-----EVPDG-YIIGNYTSSL-----198
Db	901	RDGREVEGTRTPGWTTPCLSTVAPPIHQGPAVTAVIPLPPTGTHLLFAQTGKIERLPLE 960
QY	199	GSQRYACREGFSPFVSDTSSCTGLGTWEAPKLHCQEINGN-----PPEMRHAI 249
Db	961	GNTMKTEAKAFLHVPKVI---IGLA-----FDCVDKXVWTDITEPSIGRAS 1006
QY	250	LVGNHSSRLGGVARYVCOGFSPGGKITSCTEKGTWRESTLTCTEILTAKINDVSLFND 309
Db	1007	L---H---GGEPTTIIQDLGSPGIAVDILGRNIFWTDNSLDRIE-----VAKLDG 1052
QY	310	TCVRWQINSRRINPKISVISIKQRLDPM:SVREETVNL-TTD--SRTPEVCLALYPGT 366
Db	1053	TQRRVLFETDLNPRGIVTDSVRG-----NLYTWDNRDNPKIETSYMDGT 1098
QY	367	NYTVNISTAPRRSMPAVIGFQTAEDLLEIDGSENFISIFNETCLKNRNRKRVGSEHM- 425
Db	1099	NRRILVQ---DDLGLPGLTFDAPFSQLCW/DAGTN---RACCLNPSQPSRRKALEGLQ 1151
QY	426	YQFTVLGQRWYLANFSHATSFNFTTREQVP/VCLDL-----YPTTDYTVNVTL 473
Db	1152	YPPFAVT-----SYGNLYFTDWKNS/VALDLAISKETDAFQPHKQTRLYGITALL 1202
QY	474	LRSPKRHS-----VQITIATP 489
Db	1203	SQCPQHNYCSVNGGCTHLCUATP 1227

Search completed: February 6, 2004, 17:39:37  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2004, 17:36:49 ; Search time 22 Seconds

1436.646 Million cell updates/sec  
(without alignments)

Title: US-10-066-269-58

Perfect score: 4029

Sequence: 1 MGRGFWDAGSRRLPLLL.....VGLGLAWIILFLSFAV 747

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	284	7.0	835	4	US-09-284-819-6
2	284	7.0	835	4	US-09-262-537-12
3	274.5	6.8	886	3	US-09-110-115-3
4	274.5	6.8	1833	3	US-08-479-723B-2
5	274.5	6.8	1833	5	PCT-US95-02251-18
6	260.5	6.5	1394	6	5177137-30
7	236	5.9	387	2	US-08-884-072-5
8	236	5.9	387	2	US-08-833-963C-9
9	236	5.9	387	3	US-08-980-511-3
10	236	5.9	387	4	US-09-212-163-5
11	236	5.9	1251	5	PCT-US95-02251-3
12	236	5.9	1252	1	US-08-199-780-3
13	236	5.9	1252	2	US-08-316-650-3
14	236	5.9	1253	3	US-08-479-723B-4
15	234	5.8	2199	4	US-08-793-271C-2
16	234	5.8	2199	5	PCT-US95-11634-2
17	231.5	5.7	2321	4	US-09-230-653-2
18	231	5.7	1964	4	US-09-467-997-1
19	230	5.7	2223	1	US-08-185-433-18
20	230	5.7	2523	4	US-08-899-233-3
21	228.5	5.7	816	2	US-08-820-171A-37
22	228.5	5.7	816	3	US-09-055-693-37
23	228.5	5.7	816	3	US-09-273-563-37
24	228.5	5.7	816	4	US-09-565-533-37
25	228.5	5.7	816	4	US-09-661-468-37
26	228.5	5.7	816	4	US-09-976-163-37
27	222	5.5	2703	1	US-08-185-433-19

28 222 5.5 2703 4 US-08-899-232-4 Sequence 4, Appli  
29 219 5.4 2556 1 US-08-185-432-17 Sequence 17, Appli  
30 219 5.4 2556 4 US-08-899-232-2 Sequence 2, Appli  
31 218 5.4 652 6 5258288-4 Patent No. 5258288  
32 217.5 5.4 1810 4 US-08-793-273C-4 Sequence 4, Appli  
33 217.5 5.4 1810 5 PCT-US95-11684-4 Sequence 4, Appli  
34 216 5.4 2556 1 US-08-083-590A-20 Sequence 20, Appli  
35 216 5.4 2556 3 US-08-532-384-20 Sequence 20, Appli  
36 215 5.3 635 1 US-07-907-190-1 Sequence 1, Appli  
37 215 5.3 635 1 US-07-985-691-2 Sequence 2, Appli  
38 215 5.3 635 1 US-08-436-804-2 Sequence 2, Appli  
39 215 5.3 635 1 US-08-267-387-2 Sequence 2, Appli  
40 215 5.3 676 1 US-08-282-141-4 Sequence 2, Appli  
41 215 5.3 676 1 US-08-435-434-3 Sequence 3, Appli  
42 215 5.3 676 1 US-08-435-436-3 Sequence 3, Appli  
43 215 5.3 676 2 US-08-438-863-3 Sequence 3, Appli  
44 215 5.3 676 2 US-08-438-864-3 Sequence 3, Appli  
45 215 5.3 676 3 US-08-438-862-3 Sequence 3, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-284-819-6

; Sequence 6, Application US/09284819

; Patent No. 6365712

; GENERAL INFORMATION:

; APPLICANT: Kelly, Kathleen

; APPLICANT: The Government of the United States of America

; APPLICANT: as represented by The Secretary of the

; APPLICANT: Department of Health and Human Services

; TITLE OF INVENTION: Methods and Compositions for Inhibiting Inflammation

; TITLE OF INVENTION: and Angiogenesis Comprising a Mammalian CD97 Alpha

; TITLE OF INVENTION: Subunit

; FILE REFERENCE: 015280-263100US

; CURRENT APPLICATION NUMBER: US/09/284,819

; CURRENT FILING DATE: 1999-08-20

; EARLIER APPLICATION NUMBER: US 60/027,871

; EARLIER FILING DATE: 1996-10-25

; EARLIER APPLICATION NUMBER: WO PCT/US97/19772

; EARLIER FILING DATE: 1997-10-24

; NUMBER OF SEQ ID NOS: 21

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 6

; LENGTH: 835

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: CD97amino acid sequence encoded by full-length

; OTHER INFORMATION: clone PAT276

US-09-284-819-6

Query Match 7.0%; Score 284; DB 4; Length 835;

Best Local Similarity 21.8%; Pred. No. 2,2e+15;

Matches 168; Conservative 88; Mismatches 243; Indels 270; Gaps 40;

Qy 29 GAFPGDLVCAVCHHATCQREGKKICICNVGP-----VNGRTQCVKNECQFGAT 82

Db 17 GAETQDSRGRCARWCPQNSCWATA---CRCNPGFSFSBEIITPTTETCDINECATPSK 73

Qy 83 LVCGNHTSCHTGGFYCICLEGYRATNNKTFIPNDGTFCTDIDEVTS-GLCRHGRC 141

Db 74 VSCGKFSDCWNTGSDVCVSPGYEPVSGTKTKNSENCTQDVDECCQNPRLCKSVGTC 133

Qy 142 VNTGHSPECYCMGDIPLRNGPEPFHPTDATSCTEIDCGTPPPVPGVYIIGNVTSSIGS- 200

Db 134 VNTLGSVTCQLPGF--KFIPDPKVCTDNECTS---GQNPCHSSHTCLNN-----VGSY 184

Qy 201 QVRYACREGFSV-----PDTVSSCTGLGTWSPKXHCQBIN-CGNPEMHAILVGN 253

Db 185 QCR--CRPGWQIPGSPNGPNTV--CEDYDECSGQHQCDSTVCFN-----TVGS 232

Query Match	6.8%;	Score	274.5;	DB	3;	Length	886;
Best Local Similarity	22.8%;	Pred. No.	1.5e-14;				
Matches	89;	Conservative	53;	Mismatches	125;	Indels	123;
						Gaps	17

  

QY	42	CHEHATCQOREGKKICICNTGFV-NGR	-----TQCVDKNEQCFGATLVCGNHTSCHN	93
Db	41	CPAYATCTNTVDSYCTCKQGFSSNGNHFKDPGVRCRDIIDECSS-QSPQPCGPNSSCKN		99
QV	94	TPGCFYICILEGVRATNNNTKTFIPNDGTF-CTDIDECVSGLCRGGRCWNTHGTFEYCV		152



Db 100 LSGRYKSCLDGFSFGTGNWVPGKPNFSCDTINECLTRVPEHSDCVNMGSCSC 159  
Qy 153 MDGLPRNGPFPHTTDTACTE1-DCGTFPEVPD---GYIIGNYTSLSGQVRYACR 207  
Db 160 QVGFISN-----STCEDVNECADFACPEHATCNNTVGNYS-FCN 200  
Qy 208 EGFPSVEDTVSSCTGLTWESPKLHCQBIN-----CGNPEMHA1LVGNHSS 256  
Db 201 PGFES--SSGHLSCQGL-----KASCED1LSCTEMCPINSTCNTP----- 239  
Qy 257 RLGGVARYVCOEGESPGK1-----TSVCTE----- 283  
Db 240 ---GSYFCTCHPGAPSSGQLNFDDQGVCECDIDECRQDPSTCGPNSICINALGSYSCG 296  
Qy 284 -----KGTWRESTITTEILTKINDVSLFNDTCVRWQINSRRNP-----KISYV 328  
Db 297 IVGFHPNPEGSQKDNFSCORVLFKCKEDVIPDNKQIQOCQEGTAVKPAYVSCAQNNI 356  
Qy 329 ISIKQRLDPMESVREBEVNL--TTDSRTP 356  
Db 357 FSU-----LDKVCENKTTVSLKNTTESFVP 382

RESULT 4  
US-08-479-722B-2  
; Sequence 2, Application US/08479722B  
; Patent No. 6074840  
; GENERAL INFORMATION:  
; APPLICANT: Bonadio, Jeffrey  
; APPLICANT: Yin, Wushan  
; TITLE OF INVENTION: LATENT TGF( BINDING PROTEIN (LTBP)  
; TITLE OF INVENTION: GENES, COMPOSITIONS AND METHODS  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Williams, Morgan & Amerson  
; STREET: 7676 Hillmont, Suite 250  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77040  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/479,722B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US95/02251  
; FILING DATE: 21-FEB-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/316,650  
; FILING DATE: 30-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/199,780  
; FILING DATE: 18-FEB-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Tussey, Shelley P.M.  
; REGISTRATION NUMBER: 39,458  
; REFERENCE/DOCKET NUMBER: 4100.000500/US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (713) 934-7000  
; TELEFAX: (713) 934-7011  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1833 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-479-722B-2

Query Match 6.8%; Score 274.5; DB 3; Length 1833;  
Best Local Similarity 26.4%; Pred No. 4, 5e-14;  
Matches 81; Conservative 28; Mismatches 97; Indels 101; Gaps 16;  
Qy 31 PGPDG-----LDVC-----ATCHEHATCOOREKKIKICNYGVNGRTQCVDNKCCQFGA 81  
Db 1137 PNLGNRCEDVDECEGPQSSC-RGCECKNTEGSYQCLCHQGFQVNGTMCEDVNECV--G 1193  
Qy 82 TLVCGNHTSCHNTPGGFYICLCLEGYRATNNKTFIPNDGTCTDIDECEVSLCRHGRC 141  
Db 1194 EHCAPHGELNSLGSFFCLCAPGASAE-----GTRCQDVDECAATDPC-PGGHC 1244  
Qy 142 VNTGHSFECYCMGDLPRNGPEFFHTTDTACTEID-----CGTPPEVPD 187  
Db 1245 VNTGHSFCLCETA-----SFQSPDSGCLDIDECEDREDVPCCAWRCENSP----- 1292  
Qy 188 GYIIGNYTSLSGQVRYACREGFSVPEDTVSSCTGLTWESPKLHCQINCGNPEMHA 247  
Db 1293 ---GSYRCILD-----CQGFYVAPNG-----DCIDIDECAN 1321  
Qy 248 AILVGNHS--SRLGGVARYVCOEGFE-SPGGKITSVCTEKGTVWRESTITLTCTILTKINDV 304  
Db 1322 DTVCNHNHGFCDNTDGSFRLCDQGFETSPG-----WE-----CVD-----VNEC 1361  
Qy 305 SILFNDTC 311  
Db 1362 ELMMAVC 1368

RESULT 5  
PCT-US95-02251-18  
; Sequence 18, Application PC/TUS9502251  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE  
; TITLE OF INVENTION: CELLS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: United States of America  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
; SOFTWARE: Patent In Release #1.0, Version  
; SOFTWARE: #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02251  
; FILING DATE: CONCURRENTLY HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/316,650  
; FILING DATE: 30-SEP-1994  
; CLASSIFICATION:  
; APPLICATION NUMBER: US 08/199,780  
; FILING DATE: 18-FEB-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parker, David L.  
; REGISTRATION NUMBER: 32,165  
; REFERENCE/DOCKET NUMBER: UMIC009P--  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (713) 789-2679  
; TELEX: 79-0924  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1833 amino acids

TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 PCT-US95-02251-18

Query Match 6.8%; Score 274.5; DB 5; Length 1833;  
 Best Local Similarity 26.4%; Pred. No. 4.5e-14;  
 Matches 81; Conservative 28; Mismatches 97; Indels 101; Gaps 16;  
 QY 31 PPDG-----LDVC-----ATCHEHATCOORGGKICICNYGFVNGRTQVDKNEQFGA 81  
 DB 1137 PNLGNRCEDVDECGPOSSC-RGCECKNTGSYQCLCHQGFVNGRTQVDKNEQFGA 81  
 QY 82 TLVCGNHTSCHNTGGFYCICLEGYRATNNKTFIPNDGTFCTDIDCEVSGLCRHGGRC 141  
 DB 1194 EEHCAPHGECLNSLGSFFCLCAPGASAEG-----GTRCQDVDECATDFC-PGHC 1244  
 QY 142 VNHGSPCYCMDGVLPRNGEPHPPTDA--SCTEID-----CCTPEVPD 187  
 DB 1245 VNTGFSCLCETA-----SPQSPDSGECILDIDCEDEDVPGAWRCNSP----- 1292  
 QY 188 GYIIGNYTSLSGQVRYACREGFVSVPEDT/SSCTGLGTWSPKLCHEINCNPPMRH 247  
 DB 1293 ----GSYRCILD-----CQPGFYVANG-----DCIDIDECAN 1321  
 QY 248 ALLVGNHS--SRLGVARYVCOEGFE-SPGKITSVCTEKTWRESTLTCTEILTINDV 304  
 DB 1322 DTVCGNHGFCNDTDSFCLCDQGFETSPSQ-----WE-----CVD-----VNEC 1361  
 QY 305 SLFNDTC 311  
 DB 1362 ELMMAVC 1368

RESULT 6  
 5177197-30  
 ; Patent No. 5177197  
 ; APPLICANT: KANZAKI, TETSUO; OLOFGSON, ANDERS; MOREN, ANITA;  
 ; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,  
 ; LENA; HELDIN, CARL-HENRIK  
 ; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
 ; HUMAN TRANSFORMING GROWTH FACTOR-BETA1-BINDING PROTEIN  
 ; NUMBER OF SEQUENCES: 53  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/487,343  
 ; FILING DATE: 27-FEB-1990  
 ; SEQ ID NO:30:  
 ; LENGTH: 1394  
 5177197-30

Query Match 6.5%; Score 260.5; DB 6; Length 1394;  
 Best Local Similarity 27.6%; Pred. No. 4.7e-13;  
 Matches 82; Conservative 32; Mismatches 102; Indels 81; Gaps 16;  
 QY 34 DGLDVCATCH--EHATCOOREGKICICNYGFVNG-RTQVDKNEQFGATLVCGNHTS 90  
 DB 792 EDIDECQHRHLCAHQCRNTEGSCQVCDQYRASGLGDHCEIDNECLEDKS-VC-QRCD 849  
 QY 91 CHNTGGFYCICLEGYRATNNKTFIPNDGTFCTDIDCEVSGLCRHGGRCVNHGSPFC 150  
 DB 850 CINTAGSYDCTCPDGFG-LDDNKT-----CODINECHPLGCPQGECLNTEGSPFC 900  
 QY 151 YCMDGYLPRNGEPHPPTDTSCTEIDCGPPVPVPGYIIGNYTSLSGQVRYACREGF 210  
 DB 901 VCOQG-----PSISADGTCEDIDECNNIVCDSH--GFCNTAGSFCLCYQG- 947  
 QY 211 FSVPEDTSSCTGLGTWSPKLCHEINCNPPMRHAILVGNHSSRLGVARYVCOEGF 270  
 DB 948 FOAPQD-----GQG-----CVDV-----NECELLSG-----VGEAF 974  
 QY 271 -BSPGKITSVCTEK-----GTWRESTLTCTEI-----LTKINDVSL 306  
 DB 975 CENVEGSLVCADENQYSWPTGQCRSRSTDLDVDVDPQKEKKECYNNLNDASL 1031

RESULT 7  
 US-08-884-072-5  
 ; Sequence 5, Application US/08884072  
 ; Patent No. 5872234  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bandman, Olga  
 ; APPLICANT: Corley, Neil C.  
 ; APPLICANT: Guesler, Karl J.  
 ; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/884,072  
 ; FILING DATE: Herewith  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0333 US  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 387 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GenBank  
 ; CLONE: 458228  
 US-08-884-072-5

Query Match 5.9%; Score 236; DB 2; Length 387;  
 Best Local Similarity 23.9%; Pred. No. 8.4e-12;  
 Matches 94; Conservative 46; Mismatches 139; Indels 114; Gaps 19;  
 QY 3 RGPWD-----AGPSRRLPLLLGLARGAAGAPGD-----GLDVCA----TCHHATC 48  
 DB 35 RNPADPQIPSNFSHRI-----OCAAQVEQSEHNVQCIDICTAGTNCRADQVC 84  
 QY 49 QOREGKICICNYGFVNGRTQVDKNEQFGATLVCGNHTSCHNTGGFYCICLEGYR 108  
 DB 85 INLRGSPACQPPGYQKRG-CCVDIDEC-----TIPPYCHQRCVNTPGSFYCCQSPGFQL 139  
 QY 109 TNNKTFIPNDGTFCTDIDCEVSGLCRHGGRCVNHGSPFCYCMDGVLPRNGEPHPPT 168  
 DB 140 AANNYT-----CVDINECDASNCAQ--QCYNLGSLFCQCNQY-----ELS 180  
 QY 169 TDATSCTEID-CCTPEVPDGIIGNYTSLSGQVRYACREGFFSVPEDTSSCTGLGTW 227  
 DB 181 SDRLNCEIDDECKT-----SSYLQVQCVNPEKSCMCPQGYQVVRST----- 225  
 QY 228 ESPKLCHEIN-CGNPMEHAILVGNHSSRLGVARYVCOEGF-ESPGKITSVCTEKG 285  
 DB 226 -----CODINECTNECKEDMCNHYGGFRCPNFCQDPYILTPENR--CVCP--- 274



QY 109 TNNKTFIPNDGTFCTDIDEVSGLCRHCRCVNTHTSGFECYCMDGYLPRNGPFPHT 168  
Db 140 AANNYT-----CVDINECDASNOCAQ--QCYNLSGFIQCNOGY-----ELS 180  
QY 169 TDATSCHEID-CGTPEVPDGYIIGNYTSLSQVRYACREGFFSVPEPTVSSCTGLGTW 227  
Db 181 SDRLNCEDIDECRT-----SSYLCOYQCVNMPKFSFCMCPQGVVRSRT-----225  
QY 228 ESPKLHCQEN-CGNPPENRHALLVGNHSSNLGVARYVCOBGF-ESPGKITSVCTEKG 285  
Db 226 -----CODINECTNECREDEMCMWYHGFRCYPRNFCQDPYILTENR--CVCP--- 274  
QY 286 TWRESLTCTEI-----LTKINDVSLFNDTCVRWQINS-----318  
Db 275 ---VSNAMCRELPQSIYKYSIRSDRSVPDIFQIQTIVANTINTFRKSGNENGEP 331  
QY 319 --BRINKISYVISIKGQRLDPMESVREETVNL 349  
Db 332 YLRQTSVSAAMLVLVKS-----LSGPREHIVDL 359

RESULT 10  
US-09-212-168-5  
; Sequence 5, Application US/09212168  
; Patent No. 6303765  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/09/212,168  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/884,072  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0333 US  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 387 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 458228  
US-09-212-168-5

Query Match 5.9%; Score 236; DB 4; Length 387;  
Best Local Similarity 23.9%; Pred. No. 8, 4e-12;  
Matches 94; Conservative 46; Mismatches 139; Indels 114; Gaps 19;

QY 3 RGWD-----AGPSRELLPLLLGLARCAAGPDP-----GLDVCA-----TCHHATC 48  
Db 35 RNFADPQIPSPSHRI-----QCAAYGSEHNVCQDIDECAGTHTNCRADQVC 84  
QY 49 QOREGKKICINYGFGNGRTQVDKNECQFATLVCGNHTSCHNTPGGFYICILEGYRA 108  
Db 85 INLRGSAFCQPPGYKRGGE-QCVDIDEC-----TIPPYCHQRCVNTPGSFYQCSPGQL 139  
QY 109 TNNKTFIPNDGTFCTDIDEVSGLCRHCRCVNTHTSGFECYCMDGYLPRNGPFPHT 168  
Db 140 AANNYT-----CVDINECDASNOCAQ--QCYNLSGFIQCNOGY-----ELS 180  
QY 169 TDATSCHEID-CGTPEVPDGYIIGNYTSLSQVRYACREGFFSVPEPTVSSCTGLGTW 227  
Db 181 SDRLNCEDIDECRT-----SSYLCOYQCVNMPKFSFCMCPQGVVRSRT-----225  
QY 228 ESPKLHCQEN-CGNPPENRHALLVGNHSSNLGVARYVCOBGF-ESPGKITSVCTEKG 285  
Db 226 -----CODINECTNECREDEMCMWYHGFRCYPRNFCQDPYILTENR--CVCP--- 274  
QY 286 TWRESLTCTEI-----LTKINDVSLFNDTCVRWQINS-----318  
Db 275 ---VSNAMCRELPQSIYKYSIRSDRSVPDIFQIQTIVANTINTFRKSGNENGEP 331  
QY 319 --BRINKISYVISIKGQRLDPMESVREETVNL 349  
Db 332 YLRQTSVSAAMLVLVKS-----LSGPREHIVDL 359

RESULT 11  
PCT-US95-02251-3  
; Sequence 3, Application PC/TUS9502251  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE  
; TITLE OF INVENTION: CELLS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: United States of America  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
; SOFTWARE: PatentIn Release #1.0, Version  
; SOFTWARE: #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02251  
; FILING DATE: CONCURRENTLY HERewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/316,650  
; FILING DATE: 30-SEP-1994  
; CLASSIFICATION:  
; APPLICATION NUMBER: US 08/199,780  
; FILING DATE: 18-FEB-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parker, David L.  
; REGISTRATION NUMBER: 32,165  
; REFERENCE/DOCKET NUMBER: UMIC009P--  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (713) 789-2679  
; TELEX: 79-0924  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1251 amino acids  
; TYPE: amino acid



Patent No. 5942496  
GENERAL INFORMATION:  
APPLICANT: Bonadio, Jeffrey  
APPLICANT: Roessler, Blake J.  
APPLICANT: Goldstein, Steven A.  
APPLICANT: Lin, Wushan  
TITLE OF INVENTION: METHODS AND COMPOSITIONS  
FOR STIMULATING BONE CELLS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/316,650  
FILING DATE: 30-SEP-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/199,780  
FILING DATE: 30-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: UMIC-008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1252 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-316-650-3

Query Match 5.9%; Score 236; DB 2; Length 1252;  
Best Local Similarity 23.0%; Pred. No. 4.8e-11;  
Matches 149; Conservative 48; Mismatches 218; Indels 234; Gaps 36;

QY 31 PGPDLVCATCH-----EATCQ-----REGKICI-----CNYGF 63  
DB 572 PGPDY-----SCHNAGYRHPQRYCYDVNECEAPGPGKICGNTGSGYNCHNRY 627  
QY 64 ---VNGRTQCVDRNECQFATLVCGNHTSCHNTPGSGFYICLEGYRATNNKTFIPNDG 120  
DB 628 RLHVAGGRSCVDLNEC--AKPHLCGDGFGFNFPGHYKNCYGVRL---KASRP--- 678  
QY 121 TFCITDICEVSGICRGRCVWTHGSPFC--CMDGYLPNGPEPHPTTDTATSCIEDC 179  
DB 679 PICEDIDECRDPFC--PDGKCNKPGFKCIACQGYRSGG---GACRDVNECSE--- 730  
QY 180 QTP-----PEVPDGYIIGNYTSLSGQVRYACREGFSPVPEDTVSSCTGLGTWESPKLH 233  
DB 731 GTPCSPGWCKLPGSY-----RGTCAQGI--RTGRLSCIDVDVDCGAGKV- 774  
QY 234 COEINCGNPP--EMRHAILVGNHSS-----LGVAVYVCOE 268  
DB 775 QDGICTNTPSGFCQCLSGVHLSRDRSRCEDEDDFPAACIGGDCINTNSYRCLPL 834  
QY 269 GF-----EGPG-----GKITSVCTEKGTRWSTLTACTILT-- 299  
DB 835 GHRLVGRKCKKIDECSDQDPLCLPHACENLOGSYVVCDEGFTLTQDHGCEVEOPH 894  
QY 300 -KINDVSLFNDTCVRWQINSRRINPKLSYVISIKQRLDPMESVREETVNLATDSTRTEV 358

Db 895 HKKCYLNFDDTVF-----CDSV-----LATNTQQEC 922  
QY 359 CLALYPGTNTYNTASTAPRRSMPAVIGFQTAEVDLLEDGGS--FNISIFNETCLKLNK- 415  
Db 923 CCSLGAGWDGHCBIYPCPV-----YSSAEFHSVDPGKRLHSGQGHCELCIPAHRD 973  
QY 416 -----RSRKVGSEHMYQFTVLGQRYLANFHSATSFNPTTREQVVPVCLDLY 462  
Db 974 IDECILFGAIBCKEGKCVNSQPGYECYCKGQFYDGNL-----LECVD-- 1016  
QY 463 PTTDYTNVTLRSKPRHSV-----QITIA--TPPAVKQTISNIGFNETCL--RW---- 509  
Db 1017 -----VDECLDSNCKNGVCYNTWLPACATPPA-----EYSPAQAQCLIPERWSTPQ 1064  
QY 510 RSIKTADMEEMYLPHINGQRYQWYKFAQEMTNISSSSRDPEVCLDLRP 558  
Db 1065 RDVKACAGSEERTACWGP-W-----AGPALTFD-DCCCRQPRIGTQCRP 1107

RESULT 14  
US-08-479-722B-4  
Sequence 4, Application US/08479722B  
Patent No. 6074840  
GENERAL INFORMATION:  
APPLICANT: Bonadio, Jeffrey  
APPLICANT: Yin, Wushan  
TITLE OF INVENTION: LATENT TGF ( BINDING PROTEIN (LTBP)  
TITLE OF INVENTION: GENES, COMPOSITIONS AND METHODS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Williams, Morgan & Amerson  
STREET: 7676 Hillmont, Suite 250  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77040  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,722B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US95/02251  
FILING DATE: 21-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/316,650  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/199,780  
FILING DATE: 18-FEB-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Fussey, Shelley P.M.  
REGISTRATION NUMBER: 39,458  
REFERENCE/DOCKET NUMBER: 4100.000500/FUS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (713) 934-7000  
TELEFAX: (713) 934-7011  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1253 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-479-722B-4

Query Match 5.9%; Score 236; DB 3; Length 1253;  
Best Local Similarity 23.2%; Pred. No. 4.8e-11;  
Matches 133; Conservative 53; Mismatches 206; Indels 182; Gaps 31;

QY 27 AAGAPGDLGDLVC---ATCHEHATCQOREGKIKICI-CNYGFVNGRGTQCVKNECOFGAT 82  
Db 675 ASRPEPICEDIDECRDPSTCPD-GXENKPGSPKCIACQPGYRSGGACRDVNECSEGTP 733  
QY 83 LVCCNHTSCHNTPGGFCYICLEGVYRATNNKTFIPNDGTCTDIDICEVEVGLCHRGRCV 142  
Db 734 CSPG---WCENLPGSYACTCAQIGIRTYGLS-----CIDVDECAKVCQ-DGICT 791  
QY 143 NTHGSFECYCMGDLPRNGPEPHTTDTATSCITEID-CGTPPEVDFDGYIIGNYTSLSGSQ 201  
Db 782 NTPGSFQCQLSGY-----HLSRDRKCEDIDEC-----DPPAACIGGDCINTINGS- 827  
QY 202 VRYACRGFTSVF---EDTVSSCTGLGTWESPKLHCQECNCGNPPPMRHAILVGNHSSR 257  
Db 828 YRCLCPLGHLVGRKCKKIDEC-----QDPGL-CLPHACEN-----LQGS----- 869  
QY 258 LGGVARYVCOBGFSPGPKITSVCTEKTWRNLTCTEILT---KINDVSLFNDTCVRW 314  
Db 870 -----YVC-----VCDEGFTLTQDHGCEEVEQPHHKKCYLNFDDIVF-- 908  
QY 315 QINRRINPKISYVISIKQRLDPMSVREETVNLTTDSRTPEVCLALYEGTYTWNIST 374  
Db 909 -----CDSV-----LATNVTQOECCCSLGGAGWDHCEIYP 938  
QY 375 APPRRSPAVIGFOTAEVDLDDGS--FNLSIFNETCLKLN-----RSR 418  
Db 939 CPV-----YSSAFPHSLVPDGKRLHSGQHQCELCPAHRDIDECILFGAEICEKG 989  
QY 419 KVGSEHYQFTVLQGRWLANFHSATSNFTTREQVVPVCLDLYPTDYTVNVTLLRSPK 478  
Db 990 CVNTQPGYECYCKGQFYDGNL-----LEQVD-----VDECLDESNC 1026  
QY 479 RHSV-----QITIAIPPAVKQIISNIGFNETCL---RW-----BSIKTADMEEMYLPH 524  
Db 1027 RNVGENTRGYRCACITPPA-----EYSPAQAOLIPERWSTPQRDVKAGASEERTAC 1080  
QY 525 IWGQWYQKFAQEMFTNISSSSRDPEVCLDRP 558  
Db 1081 VWGP-W---AGPALTFD-DCCCRQPLGTQURP 1108

RESULT 15  
US-08-793-273C-2  
; Sequence 2, Application US/08793273C  
; Patent No. 6482410  
; GENERAL INFORMATION:  
; APPLICANT: Crossin, Kathryn L.  
; APPLICANT: Phillips, Greg  
; APPLICANT: Prieto, Anne L.  
; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE ATTACHMENT AND  
; TITLE OF INVENTION: NEURITE OUTGROWTH, AND METHODS OF MAKING SAME  
; FILE REFERENCE: BEC0022S  
; CURRENT APPLICATION NUMBER: US/08/793,273C  
; CURRENT FILING DATE: 1997-05-22  
; PRIOR APPLICATION NUMBER: PCT/US95/11684  
; PRIOR FILING DATE: 1995-09-14  
; PRIOR APPLICATION NUMBER: 08/308,359  
; PRIOR FILING DATE: 1994-09-16  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 2199  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-08-793-273C-2

Query Match 5.8%; Score 234; DB 4; Length 2199;  
Best Local Similarity 19.8%; Pred. No. 1.6e-10;  
Matches 152; Conservative 74; Mismatches 217; Indels 326; Gaps 33;  
QY 29 GAGPFD-GLDVCA-TCHEHATCQOREGKIKICI-NYGFVNGRGTQCVKNECOFGAT 82

Db 334 GTGEDCGKPTCPHACHTQGRCE--EQG--CVCDEGPAVCDCSEKRCPCADCHNRGRVCVDG 389  
QY 74 KNECOFGAT-----LVCCNHTSCHNTPGGFCYICLEGVYRATNNKTFIPND----- 119  
Db 390 RCECDGFTGADCGKLGKCPNGSGHGRGVNGCQVCDGEGYTGDCSLQRCPCNDCHSRGRCV 449  
QY 120 -----GTFCTDID-----EVEVSL 134  
Db 450 ESKVCCEQKGYDCSDMSCPNDCHQHGRGVNGMCVDDGYTGDCRQCRQCRPCDNRGL 509  
QY 135 CRHG-----GRCVNTHGSFECYCMGDLPRNGPEPHTT 169  
Db 510 CVDGQCVCEBDGFTGPDCAELSCPNDCGGRGVNG-----QCVCHGFMGCKQCKQRCPS- 564  
QY 170 DATSCITEIDCGTPPEVDFDGYIIGNYTSLSGSQVACREGF-----FSPEDIVSSCT 222  
Db 565 -----DCHGQRCVDDGCI-----CHEGFTGLDCGQHSCESD---CN 598  
QY 223 GLGTWESPCKLHCQECNCG-----NPP-----EMR-HAILVGNHSS 256  
Db 599 NLGQCVSGRCICNCGYSGEDCSEVSPKDLVVTEVTETVNLAWDNEMRVTEYLVTYPT 658  
QY 257 RLGGVARYVCOBGFSPGPKITSVCTEKTWRNLTCTEILTCTEILTKINDVSL----- 306  
Db 659 HEGGL-----ENQFRVPDGTSTIIRELEPGVEYFIRVFAILENKKSIPIVSARVATYLP 713  
QY 307 -----FNDRCVR-----WQINSRRIN-----PKISY----- 327  
Db 714 PEGLFKSIKETSVEVWDPLDIAPETWEIIFRNKKEDEGEITKSLRRPETSRYQTGLA 773  
QY 328 -----VISIKQRLDPMSVREETVNLTTDSRTPEVCLALYPGTN 367  
Db 774 PQQVEYELSHIVKNNTRGPKLVTTTRLDAPSQIEVKDVTDTTALITWFKPLABIDGIE 833  
QY 368 YTVNISTAPPRRSPMAVIGFQTAEDVILDDGSGFNISIFNETCLKLNRSKVGSEHYQ 427  
Db 834 LTYGKDVPGDRT-----TIDLTEDESNQVSYG-----NLKPDTE--YE 869  
QY 428 FTVLQGRWLANFHSATSNFTTREQVVPVCLDLYPTDYTVNVTLLRSPKHSVQITIA 487  
Db 870 VSLISRRGDS--SNPAKETFTGLDAP-----ENLRR----- 900  
QY 488 TPPAVKQIISNIGFNETCLRWRSIKTADMEEMYLPHIHWGQRYQKEFA-----QEMT 540  
Db 901 -----VSQTDNSIT-----LEWRNGKAIDS-----YRIKYPISGGDHAEVD 938  
QY 541 FNISSSSRDPEVCLDRPCTNNVNSLRALSSEL---PVVISLTQITEP 586  
Db 939 VPKSQOATKTTLTGLRPGTEYIGVSAVKEDKESNPATINAAATELDT 987

Search completed: February 6, 2004, 17:40:10  
Job time : 24 secs